

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



**Grado en Ingeniería de Tecnologías y Servicios de
Telecomunicación**

TRABAJO FIN DE GRADO

**DETECCIÓN DE RITMO CARDÍACO MEDIANTE
ANÁLISIS DE SECUENCIAS DE VÍDEO**

**Fernando Antonio Molina Sanz
Tutor: José María Martínez Sánchez**

Septiembre 2019

DETECCIÓN DE RITMO CARDÍACO MEDIANTE ANÁLISIS DE SECUENCIAS DE VÍDEO

AUTOR: Fernando Antonio Molina Sanz
TUTOR: José María Martínez Sánchez

Video Processing and Understanding Lab
Dpto. Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Septiembre de 2019

Resumen (castellano)

El paso de la sangre a través de las venas, arterias y capilares del cuerpo genera señales invisibles para el ojo humano; sin embargo, estas señales pueden ser detectables mediante un algoritmo de procesamiento de imágenes que haga posible no solo la detección del pulso, sino también la visualización de estas señales mediante su amplificación. En la actualidad, las técnicas de detección y monitorización de ritmo cardíaco que se utilizan requieren de un contacto directo con el cuerpo, por lo que resultan incómodas para el paciente. El tratamiento de imágenes ha entrado con fuerza en el mundo de la medicina, ya que permite la detección del pulso cardíaco y la monitorización de diferentes patologías o enfermedades y, a través de esta rama de la ingeniería, pronto se podrán monitorizar pacientes a distancia. Sin embargo, la tecnología desarrollada en este campo actualmente es muy escasa.

El objetivo de este trabajo de fin de grado es la implementación de un algoritmo que, mediante el análisis de secuencias de vídeo a color, calcule el pulso cardíaco de una persona. A partir del estudio de trabajos previos, se propone un algoritmo que detecta el ritmo cardíaco de una persona mediante el análisis de secuencias de vídeo en color. Para evaluar el algoritmo, se analizan los resultados en función de la resolución, la iluminación y la distancia en una serie de vídeos que conforman un *dataset*; estos resultados se comparan con los obtenidos en los algoritmos del MIT de 2012 y de Unakafov de 2018.

Abstract (English)

The passage of blood through veins, arteries and capillaries of the body generate invisible signals to the human eye; however, this signals could be visible for an image processing algorithm that make possible not just the detection of this signals, also the visualization by amplification. Nowadays, techniques that are used for detection and monitorization of the heart rate needs direct contact with the skin, so it could be uncomfortable for the patient. Image processing has become an important field of the medicine, as it allows the pulse rate detection and the monitoring of different pathologies or diseases. Soon, we will be able to monitoring the patients remotly. However, the developed technology in this field is very scarce at the moment.

The objective of this end-of-degree work is the implementation of an algorithm that, by analyzing sequences of video in color, calculate the heart rate of a person. From the study of previous work, is proposed an algorithm that detects the heart rate of a person by analyzing the sequences of a video in color. To evaluate this algorithm, the results are evaluated based on resolution, ilumination and distance in a series of videos that make up a dataset. This results will be compared with the results obtained by the MIT 2012 and the Unakafov algorithms.

Palabras clave (castellano)

Pletismografía, foto-pletismografía remota, ritmo cardíaco, método POS, filtro *Mooving Average*.

Keywords (English)

Plethysmography, remote photo-plethysmography, heart rate, POS method, Mooving Average filter.

Agradecimientos

Este trabajo de fin de grado no hubiera sido posible hacerlo realidad sin la ayuda de mucha gente que, de alguna u otra manera, me ha apoyado y animado para conseguirlo. Quiero agradecer a los profesores de esta escuela por la formación recibida en Ingeniería de Telecomunicaciones. También, a Chema por la oportunidad que me dio cuando me propuso realizar este trabajo que recordaré durante el resto de mi vida, y por las innumerables tutorías en las que me ha guiado y en las que me ha resuelto tantas y tantas dudas.

Quiero agradecer a mi familia por todo el apoyo incondicional que me han dado, ya que ha valido para levantarme y ser más fuerte en cada momento.

Por último, quiero agradecer especialmente a Paloma su apoyo a la hora de realizar este trabajo, por haberme dejado grabarle tantos vídeos para el análisis del algoritmo y por haberme ayudado tanto a perfeccionar esta obra a la que yo con orgullo llamo mi trabajo de fin de grado.

ÍNDICE DE CONTENIDOS

1 Introducción.....	1
1.1 Motivación.....	1
1.2 Objetivos.....	1
1.3 Organización de la memoria.....	1
2 Estado del arte	3
2.1 Anatomía del corazón.....	3
2.2 Métodos tradicionales de detección de ritmo cardíaco.....	5
2.3 Detección del ritmo cardíaco mediante vídeo	6
2.3.1 Detección del ritmo cardíaco mediante movimiento.....	7
2.3.2 Detección del ritmo cardíaco mediante variaciones de color	10
2.4 Conclusiones.....	12
3 Diseño y Desarrollo	13
3.1 Introducción.....	13
3.2 Metodología.....	13
3.3 Diseño:.....	13
3.4 Desarrollo	15
3.4.1 Bloque I: Adaptación del vídeo	15
3.4.2 Bloque II: Pre-procesado de las señales raw	18
3.4.3 Bloque III: Obtención de la señal rPPG	20
3.4.4 Bloque IV: Post-procesado de la señal rPPG	21
3.4.5 Bloque V: análisis frecuencial de la señal rPPG	22
3.5 Desarrollos no incluidos:.....	23
4 Evaluación y resultados	25
4.1 Introducción.....	25
4.2 Marco de evaluación.....	25
4.2.1 <i>Dataset y Ground Thruth</i>	25
4.2.2 Métricas:	26
4.3 Pruebas y resultados:	26
4.3.1 Resolución/Iluminación.....	26
4.3.2 Distancia/Resolución	27
4.3.3 Distancia/Iluminación.....	28
4.4 Comparativa con algoritmos previos	28
4.4.1 Resolución/Iluminación.....	29
4.4.2 Distancia/Resolución	30
4.4.3 Distancia/Iluminación.....	32
4.5 Conclusiones.....	34
5 Conclusiones y trabajo futuro.....	35
5.1 Conclusiones.....	35
5.2 Trabajo futuro	35
6 Referencias	37

INDICE DE FIGURAS

FIGURA 2. 1: ANATOMÍA DEL CORAZÓN [28]	3
FIGURA 2. 2: ILUSTRACIÓN DE UN RITMO SINUSAL [4].....	4
FIGURA 2. 3: ILUSTRACIÓN DE WILLIAM HARVEY SOBRE LA CIRCULACIÓN DE LA SANGRE EN EL ANTEBRAZO [26].....	5
FIGURA 2. 4: ILUSTRACIÓN DE UN ESFIGMÓGRAFO [27]	6
FIGURA 2. 5: ILUSTRACIÓN DEL ALGORITMO EULERIAN MAGNIFICATION [8].	7
FIGURA 2. 6: ILUSTRACIÓN DEL ARCO AÓRTICO Y SUS RAMIFICACIONES [29]	8
FIGURA 2. 7: ILUSTRACIÓN DEL ALGORITMO DESARROLLADO POR EL MIT (2013) [9]	9
FIGURA 2. 8: ILUSTRACIÓN DEL ALGORITMO “PCA-BASED MAGNIFICATION METHOD FOR REVEALING SMALL SIGNALS IN VIDEO” (2013) [9]	9
FIGURA 2. 9: IMAGEN DE UNA <i>CARDIO APP</i>	10
FIGURA 2. 10: ALGORITMO EULERIAN MAGNIFICATION [8].....	11
FIGURA 2. 11: ESQUEMA DEL ALGORITMO DE MÉTODOS COMPARATIVOS REALIZADO POR UNAKAFOV (2018) [14]	12
FIGURA 3. 1: DIAGRAMA DE BLOQUES DEL ALGORITMO PROPUESTO	14
FIGURA 3. 2: SECCIÓN DEL ROSTRO DETECTADO POR VIOLA-JONES [10].....	16
FIGURA 3. 3: SELECCIÓN DE LA ROI	17
FIGURA 3. 4: ANTES (IZQDA.) Y DESPUÉS (DRCHA.) DEL REFINAMIENTO DE LA ROI	17
FIGURA 3. 5: SEÑALES RAW DESPUÉS DEL PROCESO DE “ <i>DETRENDING SIGNAL</i> ”	18
FIGURA 3. 6: FÓRMULA FILTRO MA.....	19
FIGURA 3. 7: SEÑALES RAW DESPUÉS DE APLICAR EL FILTRO MA	19
FIGURA 3. 8: SEÑALES RAW DESPUÉS DE APLICAR EL FILTRO FIR DE ORDEN 100.....	20
FIGURA 3. 9: CÁLCULO DE LA SEÑAL RPPG SEGÚN EL MÉTODO CHROM [14].....	21
FIGURA 3. 10: CÁLCULO DE LA SEÑAL RPPG SEGÚN EL MÉTODO POS [14]	21
FIGURA 3. 11: SEÑAL RPPG OBTENIDA POR EL MÉTODO POS [14].....	21
FIGURA 3. 12: SEÑAL RPPG DESPUÉS DE SUAVIZAR MEDIANTE MA.....	22

FIGURA 3. 13: SEÑAL RPPG DESPUÉS DE SUAVIZAR MEDIANTE MA.....	22
FIGURA 4. 1: REPRESENTACIÓN DEL <i>GROUND THRUTH</i>	26
FIGURA 4. 2: CÁLCULO DEL PORCENTAJE DE ERROR.	26
FIGURA 4. 3: COMPARACIÓN DEL PORCENTAJE DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA RESOLUCIÓN Y LA ILUMINACIÓN.....	30
FIGURA 4. 4: COMPARACIÓN DEL PORCENTAJE DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA RESOLUCIÓN Y LA DISTANCIA.	32
FIGURA 4. 5: COMPARACIÓN DEL PORCENTAJE DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA RESOLUCIÓN Y LA DISTANCIA.	33

INDICE DE TABLAS

TABLA 4. 1: PORCENTAJE DE ERROR EN ALGORITMO PROPUESTO EN FUNCIÓN DE LA RESOLUCIÓN Y LA ILUMINACIÓN.....	27
TABLA 4. 2: PORCENTAJE DE ERROR EN ALGORITMO PROPUESTO EN FUNCIÓN DE LA DISTANCIA Y LA RESOLUCIÓN	27
TABLA 4. 3: PORCENTAJE DE ERROR EN ALGORITMO PROPUESTO EN FUNCIÓN DE LA DISTANCIA Y LA ILUMINACIÓN.....	28
TABLA 4. 4: PORCENTAJE DE ERROR DEL ALGORITMO DE UNAKAFOV [14] EN FUNCIÓN DE LA RESOLUCIÓN Y LA ILUMINACIÓN	29
TABLA 4. 5: PORCENTAJE DE ERROR DEL ALGORITMO DEL MIT [8] EN FUNCIÓN DE LA RESOLUCIÓN Y LA ILUMINACIÓN.....	29
TABLA 4. 6: COMPARACIÓN DE PORCENTAJES DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA RESOLUCIÓN Y LA ILUMINACIÓN	29
TABLA 4. 7: PORCENTAJE DE ERROR DEL ALGORITMO DE UNAKAFOV [14] EN FUNCIÓN DE LA RESOLUCIÓN Y LA DISTANCIA.....	30
TABLA 4. 8: PORCENTAJE DE ERROR DEL ALGORITMO DEL MIT [8] EN FUNCIÓN DE LA RESOLUCIÓN Y LA DISTANCIA	31
TABLA 4. 9: COMPARATIVA DE PORCENTAJES DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA DISTANCIA Y LA RESOLUCIÓN	31
TABLA 4. 10: PORCENTAJE DE ERROR DEL ALGORITMO DE UNAKAFOV [14] EN FUNCIÓN DE LA DISTANCIA Y LA ILUMINACIÓN	32
TABLA 4. 11: PORCENTAJE DE ERROR DEL ALGORITMO DEL MIT [8] EN FUNCIÓN DE LA DISTANCIA Y LA ILUMINACIÓN.....	33

TABLA 4. 12: COMPARACIÓN DE PORCENTAJES DE ERROR DE LOS TRES ALGORITMOS EN FUNCIÓN DE LA DISTANCIA Y LA ILUMINACIÓN	33
---	----

1 Introducción

1.1 Motivación

El corazón es el órgano que bombea la sangre a todo el cuerpo; es el encargado de transportar el oxígeno y diversos materiales para alimentar las células del cuerpo. Conocer el pulso cardíaco de una persona puede ser de gran importancia a la hora de detectar posibles problemas físicos, desde cardiopatías y problemas respiratorios hasta problemas funcionales del organismo. Así pues, es de vital importancia la monitorización del ritmo cardíaco en un análisis clínico de un paciente.

Los métodos de detección y monitorización del pulso cardíaco que más se utilizan en la actualidad son métodos incómodos ya que necesitan de máquinas pesadas y parches con cables colocados en la piel del paciente, lo que hace a estos métodos muy engorrosos para el paciente. A pesar de ello, son métodos muy fiables para este fin.

Desde hace unos años, numerosos estudios han acrecentado el interés por la detección del ritmo cardíaco a través de una cámara de vídeo común o de una webcam. Estas técnicas permiten la obtención de información médica y una continua monitorización de una manera no invasiva ni obstrusiva para el paciente que un método tradicional.

Tradicionalmente, las técnicas de tratamiento de imágenes para la extracción del ritmo cardíaco se han dividido en dos vertientes: 1) la detección del ritmo cardíaco mediante el análisis de las variaciones de color que sufre la piel al paso de la sangre, y 2) la detección del ritmo cardíaco mediante el análisis de movimientos que sufre la piel o la cabeza al paso de la sangre. Ambas formas tienen como fin un único objetivo: la detección de la señal pletismográfica, que es la señal que guarda la información del ritmo cardíaco.

El objetivo de este trabajo será el de la implementación de un algoritmo que detecte de una manera muy clara esta señal mediante el análisis de imágenes en color.

1.2 Objetivos

Los objetivos que nos planteamos en este trabajo son los siguientes:

1. Examinar los algoritmos que se han desarrollado en la bibliografía para la detección del ritmo cardíaco mediante análisis de secuencias de vídeo.
2. Diseñar y desarrollar un algoritmo propio que permita la detección del ritmo cardíaco mediante análisis de secuencias de vídeo.
3. Evaluar los resultados de este algoritmo en función de la iluminación, la resolución y la distancia en una serie de vídeos que conforman un *dataset* y compararlos con los resultados de los algoritmos propuestos por el MIT [8] y por Unakafov [14].

1.3 Organización de la memoria

En este trabajo se estudiarán en primer lugar los diferentes estudios previos en la detección del ritmo cardíaco mediante análisis de secuencias de vídeo. En segundo lugar, partiendo de la base que conforman estos estudios previos, se propondrá un algoritmo que detecta el ritmo cardíaco de una persona mediante el análisis de secuencias de vídeo en color. A continuación, se evaluarán los resultados de dicho algoritmo mediante una serie de vídeos que conforman un *dataset*. La evaluación de los resultados se realizará mediante el análisis del algoritmo en función de tres parámetros de interés: resolución, iluminación y distancia. Por último, se compararán los resultados del algoritmo propuesto en este trabajo con algoritmos de estudios previos: el del MIT [8] y el de Unakafov [14].

2 Estado del arte

2.1 Anatomía del corazón

En la mayoría de organismos pluricelulares, es necesario transportar oxígeno y una serie de materiales para alimentar las células del organismo, además de recoger los desechos producidos por las propias células. Entre todos los seres vivos, la forma más común de hacer esto es mediante el torrente sanguíneo, es decir, el flujo de sangre que pasa a través del sistema circulatorio y llega a todas las células del organismo, proveyéndolas de oxígeno y de diversos materiales necesarios para su correcto funcionamiento. Por todo ello, uno de los músculos más importantes de un ser vivo es el que impulsa la sangre a través del organismo: el corazón.

En los seres humanos, el corazón está situado en el centro del pecho, entre los pulmones, levemente a la izquierda del esternón y está envuelto por una membrana llamada pericardio. Está formado por cuatro cavidades muy bien diferenciadas, dos en el lado derecho y dos en el lado izquierdo. Estas cavidades son dos aurículas en la parte superior, y dos ventrículos en la parte inferior. Las cavidades no se comunican lateralmente ya que existe un tabique separador en el medio llamado tabique interauricular. Es por esto que se suele denominar a la parte derecha *corazón derecho* y a la parte izquierda *corazón izquierdo* (Figura 2-1).

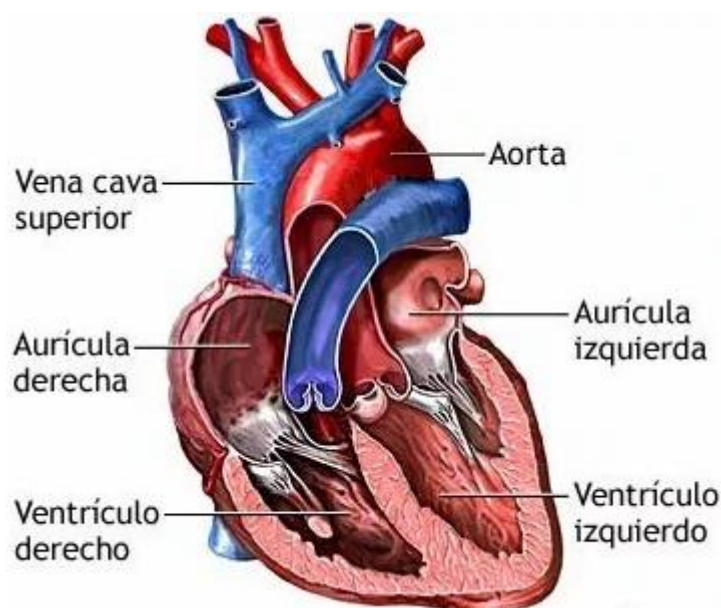


Figura 2. 1: Anatomía del corazón [28]

El corazón funciona como una bomba a presión, bombeando la sangre de las aurículas a los ventrículos y de los ventrículos a las arterias pulmonar (si la sangre viene del ventrículo derecho) o aorta (si la sangre viene del ventrículo izquierdo).

La sangre no oxigenada procedente de todo el cuerpo llega a la aurícula derecha a través de las venas cavas. La aurícula derecha está comunicada con el ventrículo derecho mediante una válvula en forma de triángulo llamada válvula tricúspide. Esta válvula se abre solo para dejar pasar la sangre de la aurícula al ventrículo cuando el corazón se relaja en un proceso llamado diástole. La sangre nunca retrocede a la aurícula. Cuando la sangre llena el ventrículo derecho, el corazón se contrae en un proceso llamado sístole. En ese

momento, la sangre es expulsada en dirección a los pulmones por medio de la arteria pulmonar. La sangre, después de oxigenarse en los pulmones, regresa al corazón por medio de las venas pulmonares y llega a la aurícula izquierda, que está conectada con el ventrículo izquierdo por medio de la válvula mitral. Esta válvula deja pasar la sangre durante la diástole al igual que la válvula tricúspide, desde la aurícula hasta el ventrículo, pero nunca al contrario. Una vez esté la sangre en el ventrículo izquierdo, el corazón se vuelve a contraer (sístole), y esta sale expulsada del corazón al resto del cuerpo a través de la arteria aorta [1].

La frecuencia a la que el corazón se contrae y se relaja marca la frecuencia cardíaca, medida en pulsaciones por minuto (ppm). Esta frecuencia varía de unas personas a otras, en función de la edad y de los hábitos físicos y saludables que tengan. Además, en una misma persona, la frecuencia cardíaca aumenta temporalmente con el ejercicio físico o ante cualquier situación de peligro. En este último caso, el cuerpo libera adrenalina, hormona que aumenta la frecuencia cardíaca para poder soportar una situación de estrés

El pulso cardíaco es la frecuencia a la que el volumen de las venas y arterias aumenta y disminuye debido al paso de la sangre a través de ellas. Esta frecuencia va acorde con la frecuencia cardíaca, y al igual que esta última, se mide en pulsaciones por minuto [2].

Mediante el electrocardiograma se obtiene la representación gráfica de la actividad eléctrica del corazón. Además, debido a la precisión que ofrece el electrocardiograma sobre la actividad del corazón, se puede sacar mucha información a través de la onda que marca el electrocardiograma. De esta manera, se puede observar gráficamente cualquier irregularidad en el ritmo cardíaco (arritmia).

La onda del electrocardiograma se puede dividir en varias ondas independientes, resultado de las contracciones y relajaciones que sufre el corazón. Estas ondas son:

- **Onda P:** Representa la despolarización de las aurículas.
- **Complejo QRS:** Conjunto de ondas que representan la despolarización de los ventrículos. Es la onda cuya amplitud positiva suele ser la más grande.
- **Onda T:** Corresponde con la repolarización de los ventrículos.
- **Onda U:** Onda de muy pequeño voltaje cuyo origen es desconocido[4].

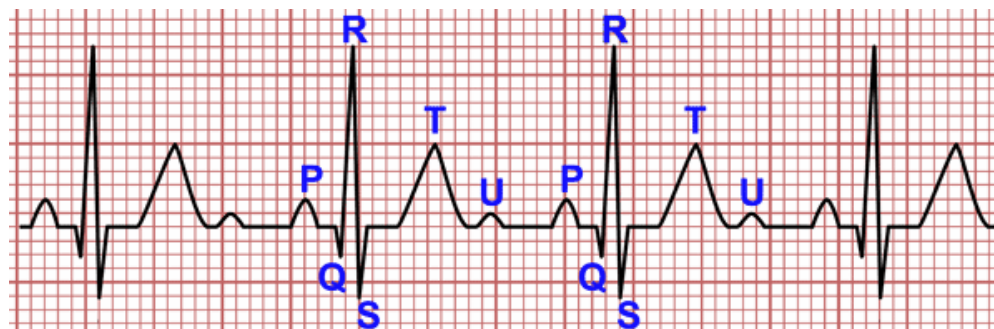


Figura 2. 2: Ilustración de un ritmo sinusal [4]

En cuanto a las patologías relativas al ritmo cardíaco, destacan las arritmias. Una arritmia es cualquier alteración del ritmo cardíaco. El ritmo normal del corazón, o también llamado ritmo sinusal, es el que tiene una frecuencia de entre 60 y 100 ppm, y en el que la onda P siempre está por delante del complejo QRS.

Las arritmias se pueden clasificar en función de la anomalía que presenten:

- **En función de la frecuencia:**
 - **Taquicardias:** frecuencia cardíaca mayor que 100ppm
 - **Braquicardias:** frecuencia cardíaca menor que 60 ppm.

- **En función del origen:**
 - **QRS estrecho:** arritmia supraventricular.
 - **QRS ancho:** arritmia supraventricular o arritmia ventricular.
- **En función del ritmo:**
 - **Regulares:** la onda P y el complejo QRS se suceden de forma regular.
 - **Irregulares:** la sucesión de las ondas es irregular [5].

2.2 Métodos tradicionales de detección de ritmo cardíaco

Durante toda la historia, numerosos científicos y médicos reconocieron la importancia de conocer el ritmo cardíaco de una persona, por lo que surgieron técnicas para medirlo.

En el siglo XVII, el médico inglés William Harvey en su libro “*Exercitatio Anatomica Motu Cordis et Sanguinis in Animalibus (Harvey, 1628)*” [26] describió por primera vez en la historia la funcionalidad del corazón y la circulación de la sangre. Este describió y estudió la posibilidad de medir el pulso cardíaco de una persona a través de su antebrazo.

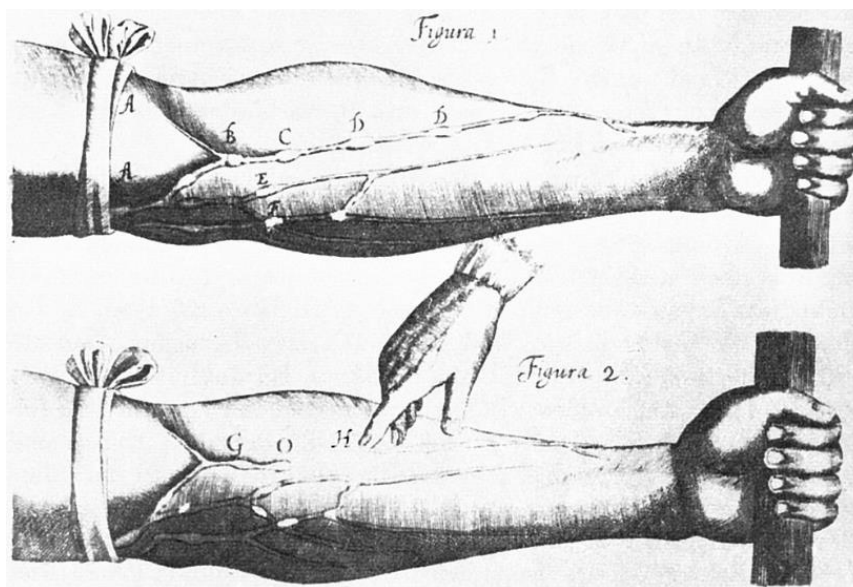


Figura 2. 3: Ilustración de William Harvey sobre la circulación de la sangre en el antebrazo [26]

Posteriormente en el siglo XIX, el físico René Laennec desarrolló el primer estetoscopio, instrumento diseñado para la auscultación de los pacientes. Este instrumento permitía conocer la frecuencia cardíaca mediante el conteo de pulsos cardíacos.

En 1854, el médico alemán Karl von Vierordt desarrolló el primer esfigmógrafo que permitía medir el pulso cardíaco y la presión sanguínea. Este método es el precursor del actual esfigmomanómetro. Posteriormente, Etienne-Jules Marey realizó modificaciones sobre la idea de Vierordt convirtiendo este sistema en portable. Este instrumento es el primero que permite registrar la onda de presión sanguínea producida por el paso de la sangre.

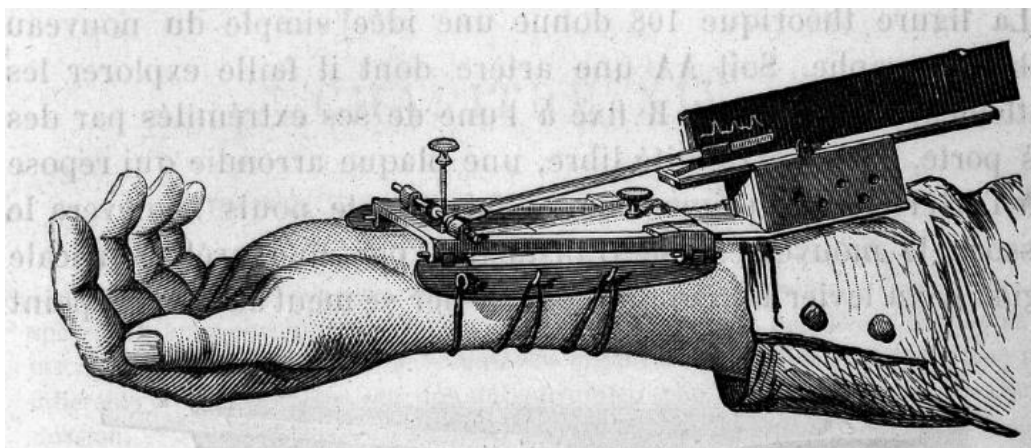


Figura 2. 4: Ilustración de un esfigmógrafo [27]

El desarrollo del electrocardiógrafo dio lugar a unos sistemas de medición más precisos. James Mackenzie, por ejemplo, inventó un polígrafo que le permitía registrar los pulsos cardíacos de una manera mucho más precisa. De esta forma, la medición del pulso cardíaco se convirtió en imprescindible en el análisis clínico [3].

En la actualidad se utiliza generalmente el electrocardiograma para medir el ritmo cardíaco de una persona en reposo. Sin embargo, este es un método incómodo para el paciente, ya que se deben colocar parches con electrodos en la piel.

2.3 Detección del ritmo cardíaco mediante vídeo

La detección de la frecuencia cardíaca mediante vídeo pretende detectar la onda que forma el pulso cardíaco mediante una cámara de vídeo, es decir, la frecuencia con la que el flujo sanguíneo llena las venas y arterias. Se basa en los principios que ofrece la pletismografía. La pletismografía es una técnica utilizada para detectar posibles problemas cardiovasculares o pulmonares. Para ello, utiliza la ley de Boyle para medir la cantidad de aire que hay en la caja torácica. Esta técnica también puede ser aplicada a las extremidades del paciente para medir la presión arterial, ya que se relaciona directamente con la presión arterial y venosa.

La foto-pletismografía (PPG) es un tipo de pletismografía que permite determinar el volumen de un tejido o de un órgano a través de un haz de luz infrarrojo. Está basada en el principio de que la sangre absorbe una cantidad detectable de luz de dicho haz, por lo que la luz reflejada variará en función de la sangre que haya. Así pues, la señal foto-pletismográfica mide las variaciones de volumen que se producen en las arterias y venas debido al ciclo cardíaco [7], y por lo tanto, se tendrá una información exacta del pulso cardíaco.

Los fotopletismógrafos más comunes son los utilizados en los *smartwatches* y, aunque es un método mucho menos incómodo que el electrocardiograma, sigue requiriendo de un contacto físico para poder medir el pulso.

Durante los últimos años, se han realizado numerosas investigaciones [6][8][14][15][24] con el objetivo de describir técnicas de medición del pulso a partir de una cámara de vídeo de bajo coste, como una *webcam* integrada en un portátil o la cámara de un teléfono móvil. Estas técnicas permiten la detección del pulso a distancia usando los principios de la pletismografía, pero todavía no se ha consensuado una denominación para estas técnicas. Durante la investigación de este trabajo se han llegado a visualizar hasta 10 nombres diferentes de esta disciplina. Sin embargo, dos de ellos parecen ser los más utilizados por

la comunidad científica: iPPG (*imaging photoplethysmography*) [14] y rPPG (*remote photoplethysmography*) [6] [15]. Además, estas técnicas demuestran que no es necesario poseer de una buena cámara de vídeo para llevarlas a cabo, sino que basta con una simple cámara de bajo coste, como una *webcam* integrada en un portátil o la cámara de un teléfono móvil.

Dos grandes enfoques han surgido con el fin de detectar el pulso cardíaco mediante vídeo: 1) la detección del ritmo cardíaco mediante el análisis de las variaciones de color que sufre la piel al paso de la sangre y 2) la detección del ritmo cardíaco mediante el análisis de movimientos que sufre la piel o la cabeza al paso de la sangre. Estas dos ramas de la investigación tienen como fin detectar la señal rPPG, pues es la que lleva la información del pulso cardíaco.

En 2012, el Instituto Tecnológico de Massachusetts (MIT) [8] desarrolló un método que permite desenmascarar señales minúsculas imperceptibles por el ojo humano. Estas señales pueden ser movimientos (por ejemplo, el movimiento que sufre una grúa a causa del viento o el latido del corazón de un recién nacido) o también minúsculas variaciones de color, como la que sufre la cara de una persona al paso de la sangre. El método fue llamado *Eulerian Magnification*. El algoritmo desarrollado por el MIT tiene como entrada un vídeo al que se le aplica una descomposición espacial, creando por cada *frame* una pirámide laplaciana o gaussiana. Después, para cada nivel de la pirámide, se aplica un filtro temporal paso banda para eliminar todas las variaciones de frecuencias consideradas como ruido. El vídeo resultante filtrado es amplificado por un factor α y añadido al vídeo original sin filtrar. Por último, se reconstruye la pirámide para obtener el vídeo final.

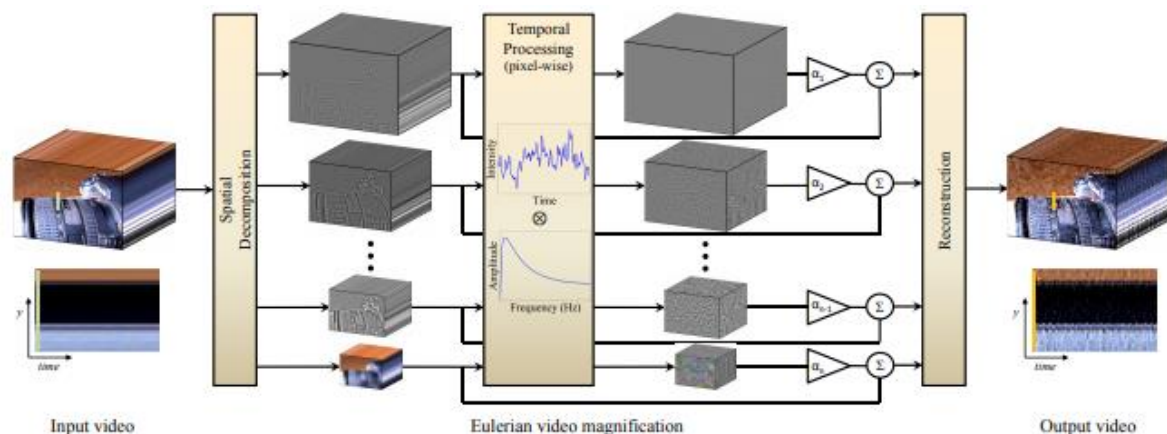


Figura 2. 5: Ilustración del algoritmo Eulerian Magnification [8].

Numerosos estudios posteriores [14] [15] [24] se han basado este algoritmo desarrollado por el MIT, tanto para análisis del pulso cardíaco por color o por movimiento.

2.3.1 Detección del ritmo cardíaco mediante movimiento

La detección de la señal rPPG a partir del movimiento de un paciente en un vídeo es una técnica que detecta los minúsculos movimientos producidos en el cuerpo como consecuencia del paso de la sangre. En el ciclo cardíaco, la sangre sale disparada del corazón desde los ventrículos al resto del cuerpo durante la diástole. La sangre pasa del ventrículo izquierdo al arco aórtico, que se ramifica, pasando después a las arterias carótidas que llegan hasta la cabeza.

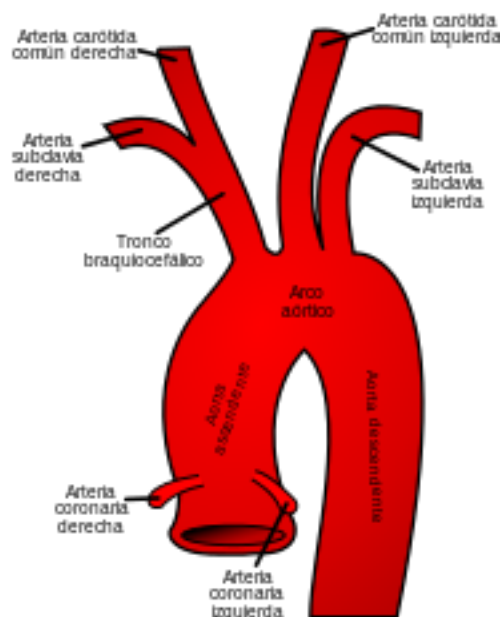


Figura 2. 6: Ilustración del arco aórtico y sus ramificaciones [29]

La sangre llega a la cabeza con un gran impulso provocando pequeños movimientos, invisibles para el ojo humano, pero detectables utilizando un algoritmo de análisis de movimiento.

En el año 2013, el instituto Tecnológico de Massachusetts (MIT) [9] estudió la posibilidad de detectar el pulso cardíaco mediante los movimientos producidos por la cabeza con el paso de la sangre. El algoritmo recibe un vídeo como entrada. Mediante el algoritmo de Viola-Jones [10] se realiza una detección facial, y por cada *frame*, se recorta esta misma región. El vídeo recortado se utiliza para el análisis del pulso cardíaco. De la región facial se extrae una (o varias) región de interés (ROI) y esta será la única región sobre la que se realizará el análisis. En concreto, el MIT [9] propone utilizar la región de la frente y la región del maxilar superior para hacer el análisis. Posteriormente, para cada ROI se realizará una detección de puntos característicos mediante el algoritmo de *Good Features to Track* [11]. Tras esto, se realiza un seguimiento de cada uno de estos puntos mediante el algoritmo KLT [12], lo que dará como resultado una serie de vectores de posición por cada punto. Por último, se aplicará el análisis de componentes principales (PCA) para obtener la trayectoria que posee mayor componente de periodicidad.

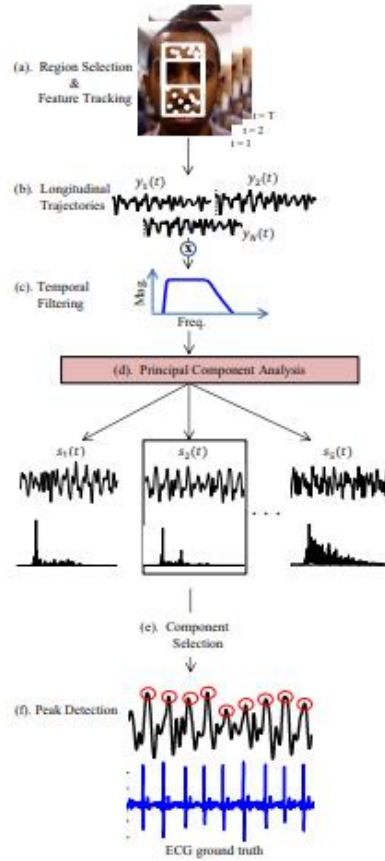


Figura 2. 7: Ilustración del algoritmo desarrollado por el MIT (2013) [9]

En 2018, la revista *Springer Nature* publicó un artículo [13] en el que optimiza el algoritmo de *Eulerian Magnification* desarrollado por el MIT [8] aplicando el análisis de componentes principales (PCA) a cada nivel de la pirámide antes de realizar la amplificación con el parámetro α , en lugar de hacer un filtrado temporal. Con esto, se obtiene por cada *frame* las componentes del vídeo cuyas variaciones espaciales coinciden mejor con las variaciones de interés; en este caso, variaciones espaciales provocadas por el flujo de la sangre. De esta manera, se consigue separar la señal buscada de todo el ruido que le rodea, consiguiendo una SNR óptima.

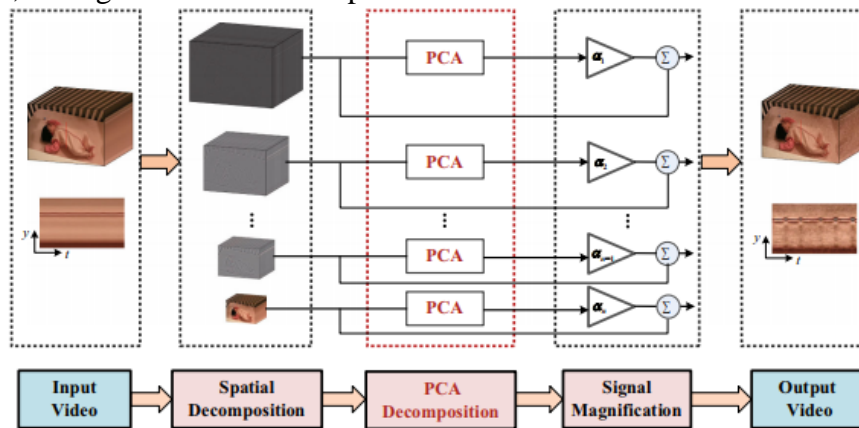


Figura 2. 8: Ilustración del algoritmo “PCA-based magnification method for revealing small signals in video” (2013) [9]

2.3.2 Detección del ritmo cardíaco mediante variaciones de color

La detección del pulso cardíaco mediante el análisis de vídeo en color es la técnica más utilizada y estudiada en este ámbito. Existen numerosas aplicaciones para móviles que son capaces de detectar el pulso cardíaco de una persona midiendo los cambios de color que se producen en uno de sus dedos. El usuario tiene que pegar la yema de uno de sus dedos a la cámara de su teléfono móvil con la luz del *flash* activada. De esta manera, el software amplifica y mide las variaciones de color provocadas por el flujo de la sangre calculando así el pulso cardíaco.

Este tipo de aplicaciones, a pesar de ser bastante precisas, siguen precisando el contacto de un aparato con la piel para medir el pulso, por lo que no pueden ser consideradas como rPPG.

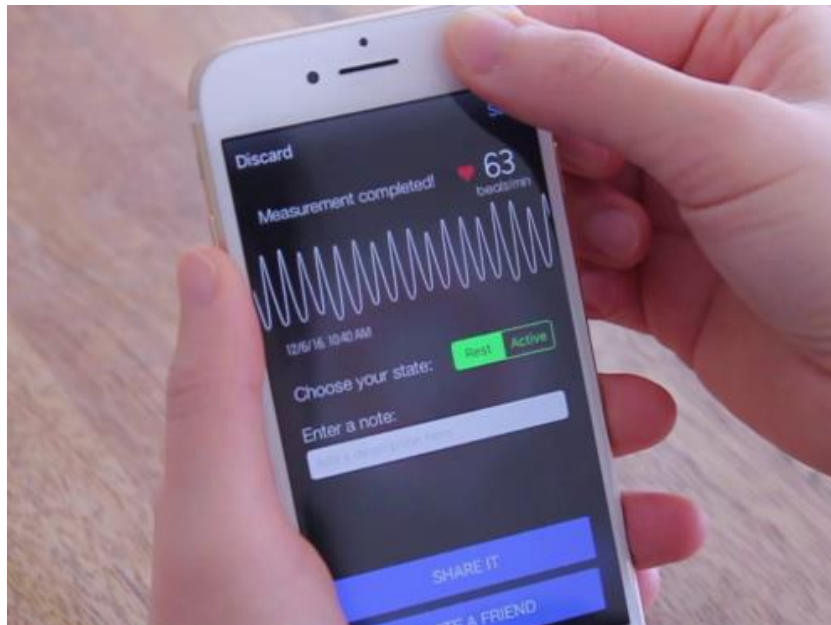


Figura 2. 9: Imagen de una *cardio app*

En 2008, la publicación del artículo *Remote plethysmographic imaging using ambient light* [6] supuso un antes y un después en esta disciplina. Por primera vez se propone la detección del pulso cardíaco utilizando una cámara convencional de bajo coste. En concreto, el algoritmo propuesto por Verkrusysse et al. [6] utilizaba el canal G como la señal rPPG basándose en que este canal es el que más energía posee y en que la hemoglobina de la sangre absorbe mejor los tonos verdes que cualquier otro color. El algoritmo tiene como entrada un vídeo en color de una resolución entre 640×480 y 320×240 . Por cada *frame*, se define una o varias regiones de interés y se separan del resto del vídeo. Después, se extrae solo la componente verde y se hace la media de cada uno de los frames, obteniéndose así la señal rPPG. El análisis frecuencial mediante la FFT da como resultado el pulso cardíaco. En este artículo también se estudian las posibles regiones de interés de la cara llegando a la conclusión de que la frente es la que tiene un mayor SNR.

Posteriormente, en 2010, Poh et al. [16] perfeccionaron el algoritmo de Verkrusysse et al. [6]. Este algoritmo usa un detector facial mediante el algoritmo Viola-Jones [10] en cada *frame* para monitorizar mejor a la persona y permitir el movimiento de esta. Además, utiliza los tres canales y no solo el verde para obtener la señal rPPG. Esto lo hacía

mediante un análisis de componentes independientes (ICA). Asimismo, estos autores proponen tomar como señal rPPG la segunda componente del resultado del ICA, ya que esta señal es la que más componentes periódicas posee. Después, realiza un análisis espectral mediante FFT.

En 2012, el MIT publica su artículo de *Eulerian Magnification* [8], en el cual proponen una manera de visualizar la señal rPPG. Utilizan una descomposición espacial (mediante pirámides Laplacianas o Gaussianas) para, mediante un filtrado temporal, amplificar la señal de interés mediante un parámetro α . Esta señal de interés se aplica posteriormente al vídeo original y, de esta manera, se visualiza de forma óptima la variación de color que sufre la piel debido al paso de la sangre. La señal rPPG se calculará haciendo la media de cada uno de los frames de los tres canales; por lo tanto, habrá tres señales rPPG. El pulso cardíaco se calcula mediante un análisis espectral de estas tres señales.

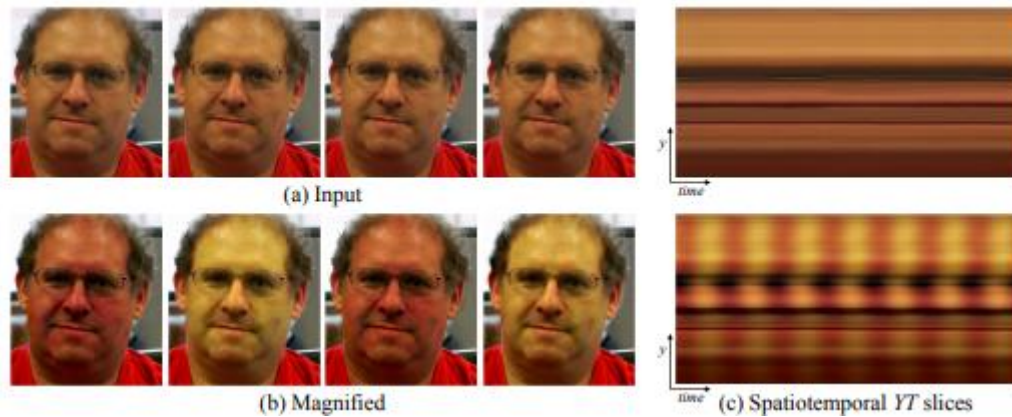


Figura 2. 10: Algoritmo Eulerian Magnification [8]

En los últimos cinco años se han realizado numerosos estudios con el objetivo de adquirir, con la mejor SNR posible, la señal rPPG. Para conseguirlo, dichos estudios tenían los siguientes objetivos: la detección de la región facial de interés que diese los mejores resultados posibles para distintos usuarios [17] [18] [19]; la utilización de diferentes tipos de filtrado para optimizar la eliminación de ruido, como los filtros *Moving average* [20]; el uso de diversos métodos para obtener la señal rPPG a través de los tres canales de color (R, G y B) utilizando PCA o ICA [21]; y el estudio de los diversos tipos de análisis espectral, como la FFT [6][8][21], o la detección de picos de la señal pletismográfica [22].

En 2018, Anton M. Unakafov [14] comparó los mejores métodos de detección de ritmo cardíaco mediante análisis de vídeo en color y los métodos o sub-algoritmos que dieron mejores resultados en cada parte del proceso de detección fueron incluidos en el algoritmo final.

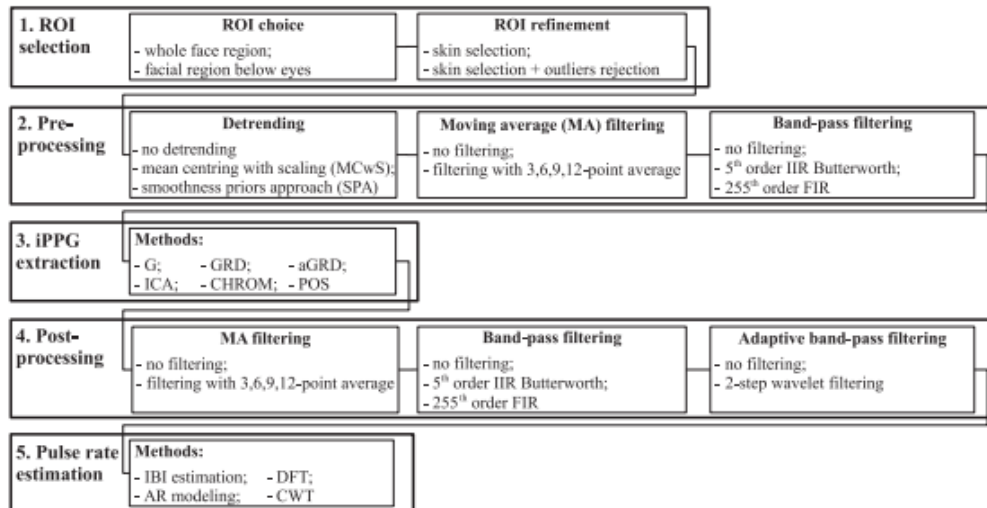


Figura 2. 11: Esquema del algoritmo de métodos comparativos realizado por Unakafov (2018) [14]

Unakafov divide los algoritmos de detección de ritmo cardíaco en cinco bloques. En cada uno de ellos, compara los sub-algoritmos que hayan sido utilizados hasta la fecha. Estos bloques son:

- **Detección de ROI:** se encarga de extraer la ROI y separarla del resto del vídeo.
- **Pre-procesado:** obtiene las señales correspondientes a cada color: R, G y B. Tras ello, elimina la tendencia que puedan tener y realiza filtrados frecuenciales para eliminar el ruido en las frecuencias que estén fuera del rango de interés.
- **Extracción de la señal iPPG/rPPG:** se calcula la señal foto-pletismográfica a partir de las señales R, G y B.
- **Post-Procesado:** El post-procesado se realiza para asegurarse de que la señal foto-pletismográfica posee una buena calidad y que no posee ruido o armónicos que molesten en el resultado final.
- **Estimación del pulso cardíaco:** realiza un análisis espectral de la señal foto-pletismográfica para calcular el pulso cardíaco final.

Finalmente, Unakafov propuso una serie de sub-algoritmos que habían dado los mejores resultados en cada bloque (Figura 2.11). De esta manera, facilitó las investigaciones futuras.

2.4 Conclusiones

Como se ha mostrado en este capítulo, los estudios de los últimos años sobre el tema han ido enfocados a la detección de la señal rPPG, de forma que esta sea lo más clara y robusta posible. Por lo tanto, se ha decidido profundizar en la detección del pulso cardíaco mediante el análisis de vídeo en color. Para ello, se desarrollará un algoritmo capaz de realizar un reconocimiento facial, obtener la señal foto-pletismográfica y optimizar dicha señal para aumentar la SNR y conseguir así un resultado óptimo.

La obtención del ritmo cardíaco en color es un método muy poco sensible al movimiento del sujeto analizado, al contrario que la detección por análisis de movimiento, ya que cualquier movimiento del sujeto puede alterar drásticamente el resultado. Sin embargo, la detección con color es muy sensible a cambios de iluminación, por lo que se estudiarán en este trabajo las mejores condiciones lumínicas para el análisis.

3 Diseño y Desarrollo

3.1 Introducción

En el presente trabajo, se ha decidido desarrollar un algoritmo capaz de detectar el pulso cardíaco mediante el análisis de secuencias de imágenes en color. El algoritmo se basará en las propuestas realizadas por Unakafov [14]. El objetivo principal del algoritmo propuesto es el de obtener la señal foto-pletismográfica, es decir, aquella que se extrae mediante el análisis de las pequeñas variaciones de color que sufre la piel de una persona como consecuencia de la absorción de la luz que realiza la sangre. Dicha señal debe ser desenmascarada de todo el ruido que pueda tener, haciendo que esta sea una señal robusta. También se tendrá como objetivo que dicho algoritmo tenga un bajo coste computacional y que su sensibilidad a la iluminación, resolución y distancia de análisis sea la mínima posible.

Por lo tanto, en este apartado se describirá de manera detallada el diseño del algoritmo propuesto. Después se desarrollarán cada una de las partes de este algoritmo y se describirán aquellos desarrollos que no han sido incluidos finalmente en el algoritmo.

3.2 Metodología

La metodología que se ha seguido en este trabajo es continuista con los estudios realizados anteriormente. Verkruysse et al. [6] fueron los primeros en proponer la detección del ritmo cardíaco mediante cámaras convencionales de bajo coste y, desde entonces, numerosos estudios han continuado con esta tendencia. Además, el estudio de esta técnica utilizando material de bajo coste servirá de base a futuras aplicaciones cuyo objetivo sea la obtención del pulso cardíaco a distancia, haciendo que estas aplicaciones puedan ejecutarse utilizando cámaras convencionales como las de un teléfono móvil.

Por todo ello, para la realización de este trabajo se tomarán vídeos de un sujeto a distancias variables con una cámara Full HD 1080p (1920 x 1080), integrada en el teléfono móvil Huawei P20-lite. Los vídeos serán tomados con una configuración normal del sensor, sin ningún tipo de filtro aplicado que pueda alterar el resultado final. Tendrán una resolución máxima de 1920 x 1080 y mínima de 640 x 480 y tendrán todos ellos aproximadamente 30 frames por segundo (fps).

3.3 Diseño:

El algoritmo propuesto está basado en las recomendaciones de Unakafov [14] y se divide en cinco bloques independientes. Cada uno de estos bloques posee una función determinada, siendo posible utilizar en cada uno de ellos los sub-algoritmos que se crean convenientes para obtener el mejor resultado final posible.

El algoritmo propuesto tiene el siguiente esquema:

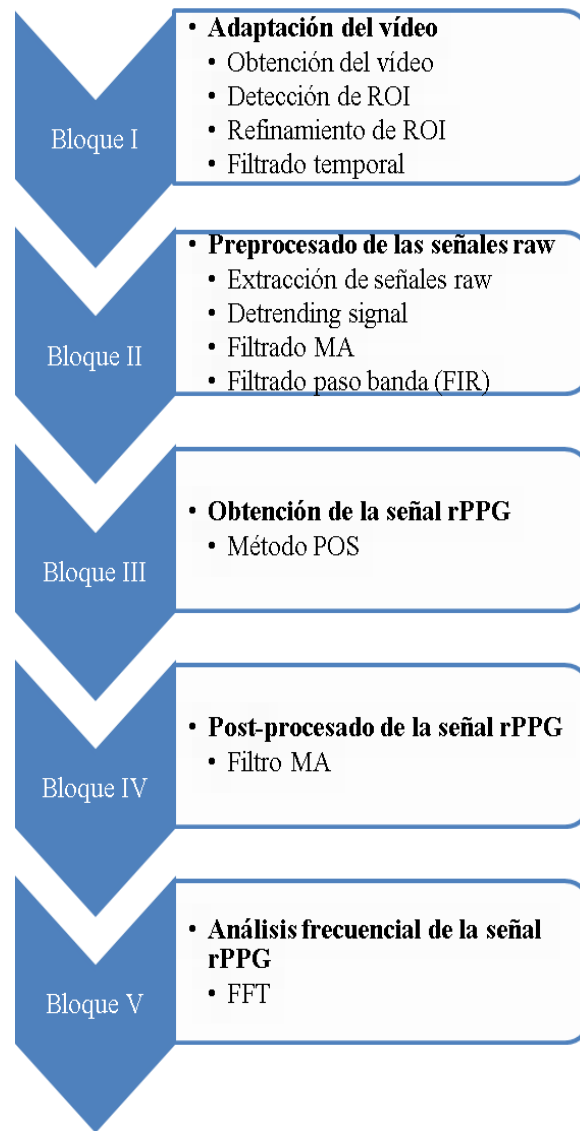


Figura 3. 1: Diagrama de bloques del algoritmo propuesto

El primer bloque es el encargado de adaptar el vídeo bajo análisis a unas condiciones óptimas. El objetivo es eliminar todos aquellos píxeles que formen ruido y que puedan alterar el resultado final. En primer lugar, se extrae el vídeo del *dataset*. Una vez cargados todos los *frames* del vídeo, se procede a detectar la región de interés o ROI. Por motivos de coste computacional, la ROI será detectada en el primer *frame* del vídeo, recortándose esta misma región en el resto de *frames* del vídeo. Para poder hacer esto, el sujeto analizado en el vídeo debe mantener una posición estática durante todo el vídeo y conseguir así que la región seleccionada en el primer *frame* sea la misma para el resto. Durante la realización del algoritmo se ha utilizado el método de selección manual de la ROI para hacer más cómoda la selección y el análisis de diferentes ROIs. En el algoritmo final se utilizará, sin embargo, la detección de la ROI automáticamente. Para ello, se realizará primero una detección facial mediante el algoritmo de Viola-Jones [10] y después se seleccionará la ROI. Una vez detectada la ROI, se procederá a recortarla en cada *frame* del vídeo. La elección de la ROI es un punto muy importante debido a que la ROI es la región sobre la que se va a aplicar el algoritmo para extraer la rPPG y, en definitiva, es la región donde se va a trabajar. En este algoritmo se usará como ROI la región de la frente. Esta región se separará del resto del vídeo y será la misma para el resto de *frames*.

Una vez seleccionada la región de interés, se procederá a eliminar todos aquellos píxeles que sean detectados como no-piel. Esto se realizará mediante un simple detector de piel utilizando el mapa de color HSV. De esta manera, todos aquellos píxeles (x,y) que caigan fuera de un rango determinado serán puestos a cero.

Por último, este bloque realizará un filtrado temporal, haciendo que todos los píxeles (x,y) que varíen con una frecuencia superior o inferior al rango de frecuencias de análisis también sean puestos a cero en todos los *frames*.

El segundo bloque tiene por objetivo obtener las señales en bruto de los colores (R, G y B) o señales *raw*. Estas señales se obtendrán mediante la media de los píxeles de cada *frame*, es decir, para la señal *raw* R se realizará la media de todos los *frames* en el canal R, y de la misma manera se obtendrán las señales *raw* G y B. Además, en cada señal *raw* se eliminará la pendiente que pueda tener y se centrará la señal en el origen haciendo que sus medias sean iguales a cero. Posteriormente, se aplicará un filtro MA (*Moving Average*) que permitirá suavizar la señal eliminando ruido de altas frecuencias. Por último, se realizará un segundo filtrado paso banda utilizando para ello un filtro FIR con orden de entre 100 y 128, como recomienda Unakafov [14].

El tercer bloque consiste en la obtención de la señal rPPG a partir de las tres señales *raw*. Para la realización de este bloque se han probado diversos métodos utilizados anteriormente por otros estudios y se ha determinado que los que dan los mejores resultados son el método ICA [14] [23], el método CHROM [14] [24] y el método POS [14] [24]. Estos métodos darán de formas diferentes la señal rPPG de la cual se extraerá el pulso final. En apartados posteriores se analizará el resultado de cada uno de estos métodos.

El siguiente bloque es el que procesa la señal rPPG. La reducción de dimensionalidad realizada en el bloque anterior calculaba a partir de las tres señales *raw* una sola señal. Esto puede dar como resultado que la señal rPPG tenga picos espurios y componentes de alta frecuencia añadidos como consecuencia de este proceso de reducción de dimensionalidad. Para suprimir estas componentes frecuenciales se procederá a suavizar la señal mediante un filtro MA. De esta manera se suprimirán estas componentes que podrían dar lugar a resultados erróneos.

El último bloque consiste en el análisis espectral de la señal rPPG para la obtención del pulso cardíaco. Dicho análisis se realizará mediante la Transformada Rápida de Fourier, que dará como resultado el pulso cardíaco medio que posee el sujeto a lo largo de todo el vídeo.

3.4 Desarrollo

El código del algoritmo propuesto se ha realizado en Matlab R2016a. El script principal es el que ejecuta el algoritmo al completo. Este script se va introduciendo en otros scripts independientes o funciones, en los que están implementados los bloques de los que se compone este algoritmo. Antes de ejecutar el algoritmo se deberá elegir un vídeo para su análisis. Los videos consisten en la grabación de un sujeto que permanece inmóvil frente a la cámara. Tienen una duración de entre 10 y 30 segundos, y todos ellos poseen una frecuencia de cuadro de alrededor de 30 fps. Una vez seleccionado el vídeo del *dataset*, se podrá ejecutar el algoritmo.

3.4.1 Bloque I: Adaptación del vídeo

Nada más comenzar la ejecución, el programa cargará el vídeo y lo almacenará en una variable del programa. El primer paso para la detección de la señal rPPG será la detección de la región de interés donde se desea analizar al sujeto. Se ha decidido utilizar la región de

la frente debido a que es la zona de la cara más plana y donde no suelen aparecer sombras o reflejos como en otras partes del rostro. Además, esta región no se ve afectada en los usuarios con barba. En definitiva, es la región que ofrece los mejores resultados. Para la detección de esta región, se realiza un reconocimiento facial con el algoritmo de Viola-Jones [10] implementado en Matlab.

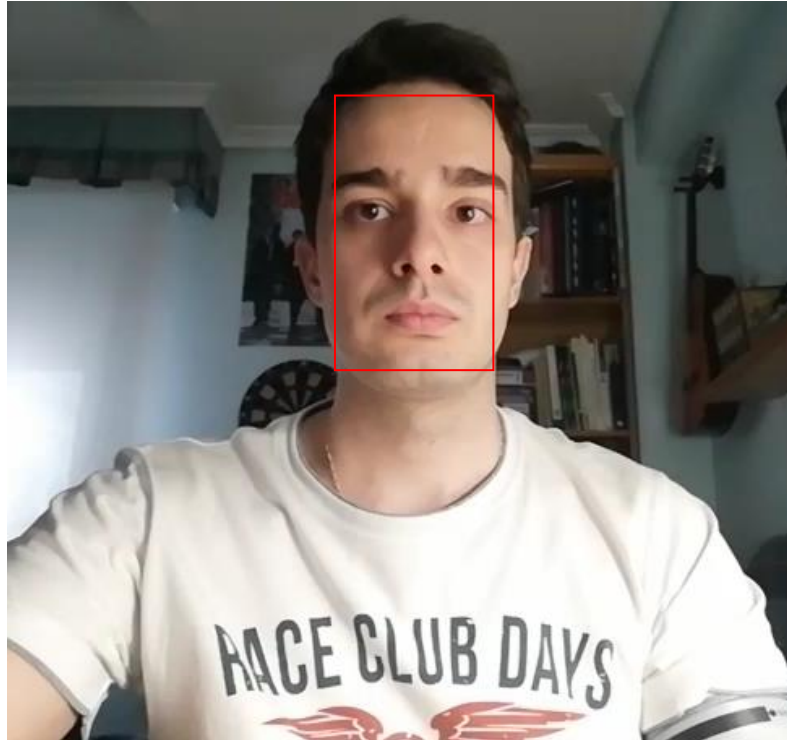


Figura 3. 2: Sección del rostro detectado por Viola-Jones [10]

Para escoger la región de la frente como la ROI, se seleccionará el rectángulo que queda al eliminar el 73 % de la parte inferior del rectángulo detectado por Viola-Jones. De esta manera, los ojos y las cejas quedan eliminados de la ROI en la mayoría de sujetos. El resultado de este proceso se muestra en la Figura 3.3.

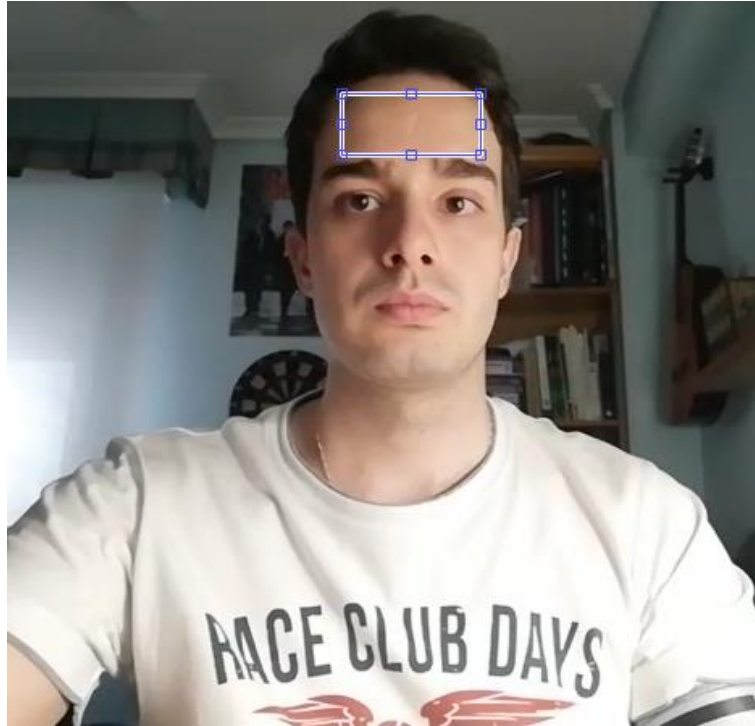


Figura 3. 3: Selección de la ROI

Una vez seleccionada la ROI, se procede a recortar en cada *frame* esta región. A partir de ahora, el vídeo sometido al análisis serán los *frames* formados por esta región de interés. La selección óptima de la ROI no supone, sin embargo, una garantía de haber descartado de la imagen píxeles que representen no-piel. Es por este motivo que se debe asegurar que todos los píxeles analizados en cada *frame* del vídeo sean piel, y de esta manera poder optimizar el resultado. Para ello, se utilizará un detector de piel basado en el espacio de color *Hue Saturation Value* (HSV). Para todos los *frames*, se descartarán todos los píxeles (x,y) que queden fuera de unos rangos determinados en el espacio de color HSV [14]. Estos rangos son:

- **Hue:** [0, 46].
- **Saturation:** [23, 132].
- **Value:** [88, 255].

Todos los píxeles detectados como no-piel serán puestos a cero.

La siguiente figura representa la región de interés antes y después de realizarse este proceso:

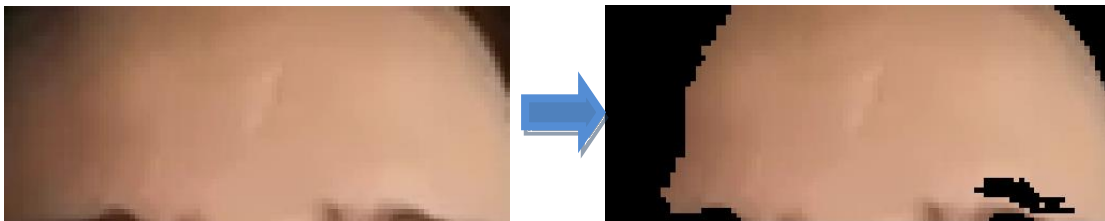


Figura 3. 4: Antes (Izqda.) y después (drcha.) del refinamiento de la ROI

Todos los píxeles en las posiciones espaciales (x,y) que queden descartados en un solo *frame* quedarán descartados en el resto de *frames*.

Por último, se realizará un filtrado paso banda temporal para descartar aquellos píxeles que varíen temporalmente con una frecuencia fuera de un umbral determinado. Para ello, se utilizará un filtro ideal paso banda con frecuencia de corte mínima de 0.833 Hz y frecuencia de corte máxima de 3 Hz. Estas frecuencias corresponden con 50 ppm y 180 pmm respectivamente. De esta forma se eliminarán aquellas frecuencias que quedan fuera de la frecuencia cardíaca normal.

3.4.2 Bloque II: Pre-procesado de las señales raw

Este bloque comienza calculando las señales *raw*, es decir, las señales que indican la variación de cada color (rojo, verde y azul) a lo largo del vídeo. Para extraer estas señales, se hará la media de cada canal de cada *frame*, quedando, por lo tanto, tres señales *raw* correspondientes a los tres canales (rojo, verde y azul).

Después, se procesarán para conseguir que la SNR final sea óptima, eliminando todo el ruido posible. Para ello, en primer lugar, se realizará un proceso llamado *detrending signal* [14]. Así, las señales *raw* tendrán media cero y se eliminará la pendiente que puedas tener. De esta forma, se consigue que la componente pulsátil de la señal rPPG sea aquella con la mayor amplitud y que el resto de componentes queden minimizadas o eliminadas. Esto puede considerarse como un filtro paso alto, ya que elimina las variaciones lentas de la señal temporal. Este proceso se realizará mediante la función de Matlab *detrend*.

Una vez realizado el proceso *detrending signal*, las señales *raw* quedan de la siguiente manera:

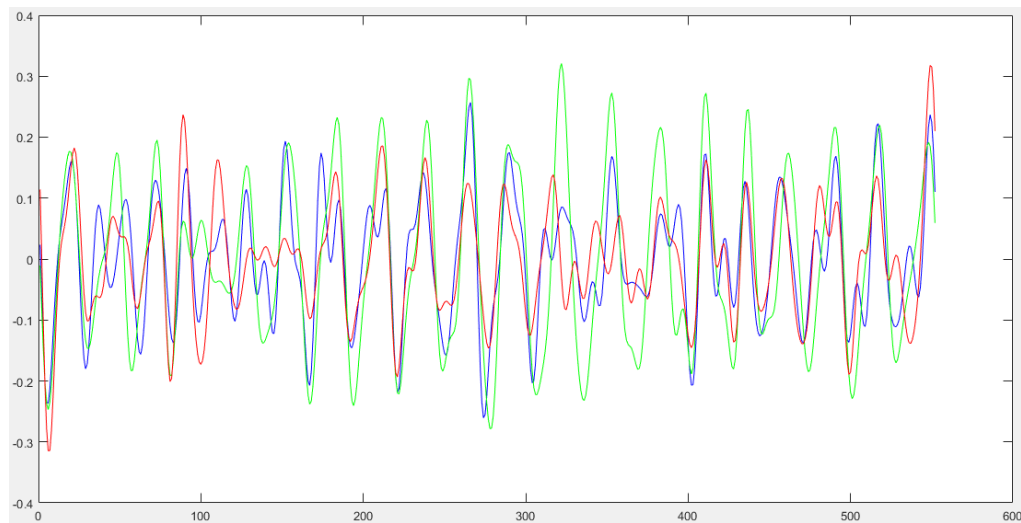


Figura 3. 5: señales raw después del proceso de “detrending signal”

Como se observa en la figura anterior, las señales *raw* están formadas por muchas componentes aleatorias que forman picos de alta frecuencia y no poseen una clara estructura periódica. Para poder extraer la señal rPPG a partir de las señales *raw*, será necesario aplicar un filtro que suavice estas señales eliminando el ruido de alta frecuencia. Para ello, se utilizará un filtro *Moving Average*. Este filtro consiste en una ventana móvil que promedia un número determinado de valores, formando de esta manera un filtro paso bajo [15]. La forma de este filtro viene representada por la siguiente fórmula:

$$c(t) = \frac{1}{M} \sum_{k=0}^{M-1} c^0(t - k)$$

Figura 3. 6: Fórmula filtro MA

Este filtro se encarga de eliminar las frecuencias superiores a $(n/M)*\text{fps}$. Unakafov [14] recomienda el uso de $M \leq (\text{fps}/4)$ debido a que la frecuencia cardíaca máxima de una persona rara vez llegará a 4 Hz. El filtro MA será aplicado a las señales *raw* utilizando la función de Matlab *movmean*. Al aplicar el filtro, las señales *raw* quedarán suavizadas, como se aprecia en la Figura 3.7:

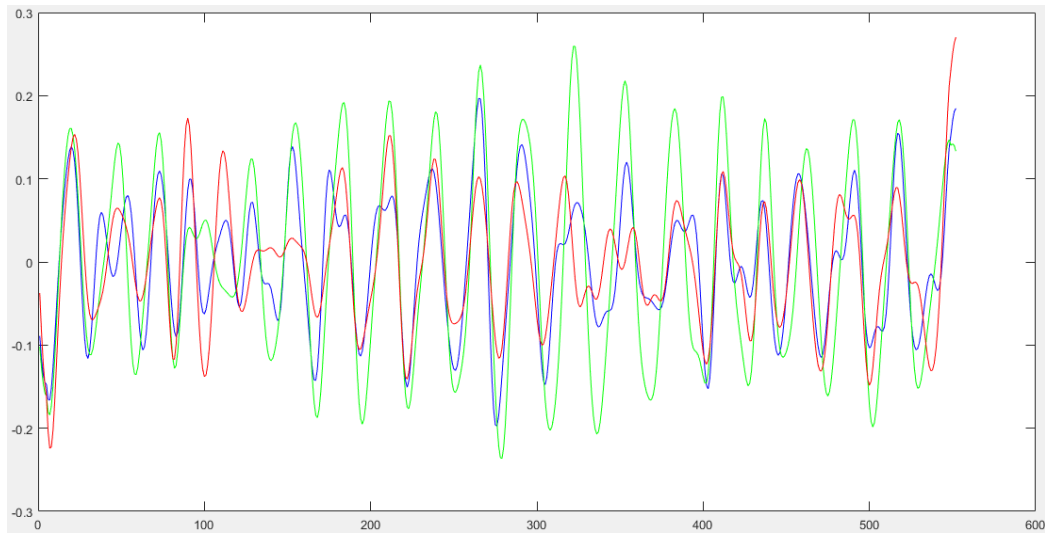


Figura 3. 7: Señales raw después de aplicar el filtro MA

Como se observa, la mayoría de los picos de alta frecuencia quedan suavizados por efecto del filtro y, de esta manera, se han eliminado las componentes espurias en frecuencias aleatorias y se ha conseguido hacer más consistente a la componente pulsátil de la señal rPPG.

Para terminar este bloque, se realizará un último filtrado paso banda para eliminar las componentes que queden fuera del rango frecuencial de estudio, y así, poder obtener la señal rPPG de una forma más clara y sin introducir componentes espurias que afecten al resultado. El filtro utilizado es un filtro FIR de orden 100, propuesto por Unakafov [14], debido a que ha dado los mejores resultados para la mayoría de sujetos analizados.

Las señales *raw*, al final de este bloque, quedarán de la siguiente manera:

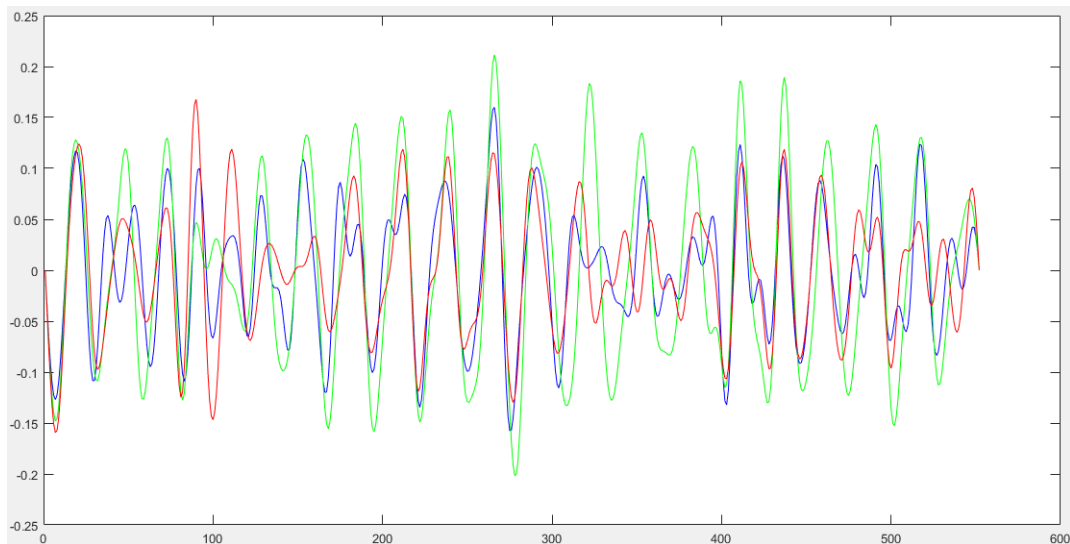


Figura 3. 8: Señales raw después de aplicar el filtro FIR de orden 100

3.4.3 Bloque III: Obtención de la señal rPPG

Este bloque será el encargado de obtener la señal rPPG a partir de las tres señales *raw*. Para ello, se han estudiado los diversos métodos que proponen Unakafov [14] y Rouast et al. [15]. La obtención de la señal rPPG depende en gran medida del entendimiento de la misma. Numerosos estudios han utilizado como señal rPPG la señal *raw* verde [6] basándose en que la hemoglobina de la sangre captura mejor las componentes verdes de la luz, además de que el verde es el color primario captado con mayor energía por las cámaras de vídeo. Sin embargo, estudios posteriores [24] [14] [15] han demostrado grandes resultados al entender la señal rPPG como una combinación de las tres señales *raw*. Es por ello que se han analizado dos de estos métodos que han dado los mejores resultados en estudios anteriores: el método ICA, y el método POS, evolución del algoritmo CHROM propuesto por Wang et al. [24].

El método ICA (*Independent Component Analysis*) es un método de *Blind Source Separation* [14] que trata de separar de un conjunto de muestras los diferentes patrones que se observan. Por ejemplo: en una fiesta donde hay música y un enorme griterío de fondo, resultaría imposible separar cada una de estas señales sonoras y escucharlas por separado; sin embargo, utilizando un análisis de componentes independientes (ICA) se conseguiría separar cada una de estas señales entremezcladas y poder escucharlas individualmente. Por lo tanto, aplicando ICA a las señales *raw* daría como resultado las componentes que forman estas señales. Entre estas componentes se encuentra la señal rPPG que, según Unakafov [14], es aquella componente que posea mayor periodicidad. Sin embargo, la señal buscada no siempre va a tener la mayor componente periódica debido al ruido que pueda haber en la imagen, por lo que se ha decidido no utilizar este método en el algoritmo final.

El otro método analizado es el método POS, propuesto por Unakafov [14] y Wang et al. [24]. Este método consiste en una evolución del método CHROM, de Wang et al. [24].

El método CHROM calcula una señal única a partir de la desviación típica de las señales *raw*. La señal rPPG utilizando CHROM será la representada en la siguiente figura:

$$iPPG^0(t) = x_1(t) - \frac{\sigma_1(t, L)}{\sigma_2(t, L)}x_2(t),$$

Figura 3. 9: Cálculo de la señal rPPG según el método CHROM [14]

Las señales x_1 y x_2 serán: $x_1(t) = 0,77r(t) - 0,51g(t)$; $x_2(t) = 0,77r(t) + 0,51g(t) - 0,77b(t)$. El método POS es una simplificación del método CHROM, donde la señal x_1 será: $x_1(t) = g(t) - b(t)$; y la señal será: $x_2(t) = g(t) + b(t) - 2r(t)$. La señal rPPG se obtendrá con el método POS utilizando la siguiente fórmula:

$$iPPG^0(t) = x_1(t) + \frac{\sigma_1(t, L)}{\sigma_2(t, L)}x_2(t)$$

Figura 3. 10: Cálculo de la señal rPPG según el método POS [14]

El método de obtención de la señal rPPG utilizado finalmente en el algoritmo ha sido el método POS. Al aplicar este método a las señales *raw* de la Figura 3.8, el resultado es la señal contenida en la siguiente figura:

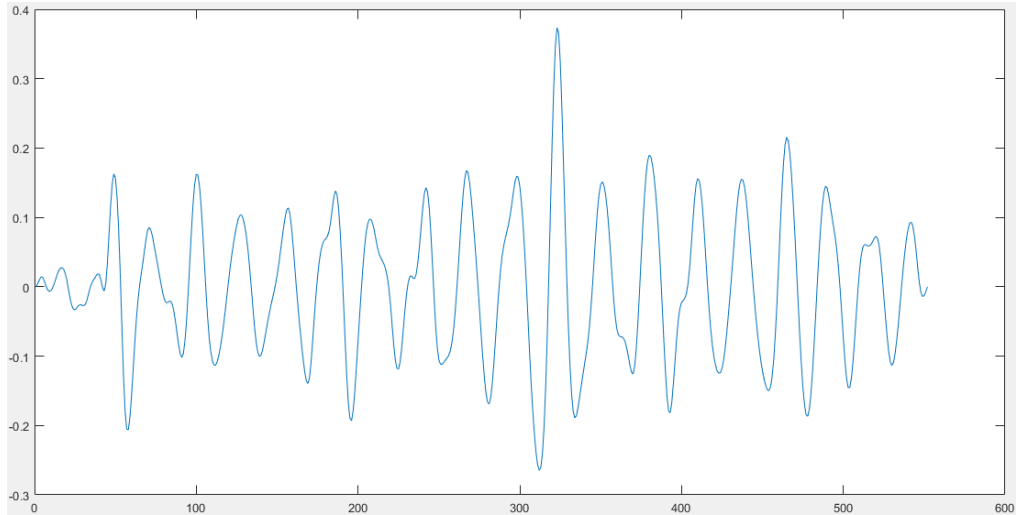


Figura 3. 11: Señal rPPG obtenida por el método POS [14]

Como se observa en la Figura 3.11, la señal rPPG obtenida del método POS está bastante bien definida. Este método tiene la ventaja de hacer más fuerte la componente común a las señales de entrada, es decir, al tener las señales *raw* formas periódicas, la señal rPPG tendrá fuertes componentes periódicas y eliminará las componentes no periódicas. Por esta razón, como se aprecia en la figura, la señal rPPG tiene una fuerte componente periódica y pocos picos espurios. Por todo ello, se ha decidido utilizar este método para obtener la señal rPPG.

3.4.4 Bloque IV: Post-procesado de la señal rPPG

A pesar de que el método POS elimina los picos espurios de las señales *raw*, al ser este método una combinación matemática de las señales de entrada, pueden aparecer a la salida componentes de alta frecuencia que pueden perturbar el resultado drásticamente. Es por ello que se ha decidido hacer un último suavizado mediante un segundo filtro MA con las

mismas características que el utilizado en el bloque II, es decir, para $M \leq (\text{fps}/4)$. El resultado final de la señal rPPG está representado en la siguiente figura.

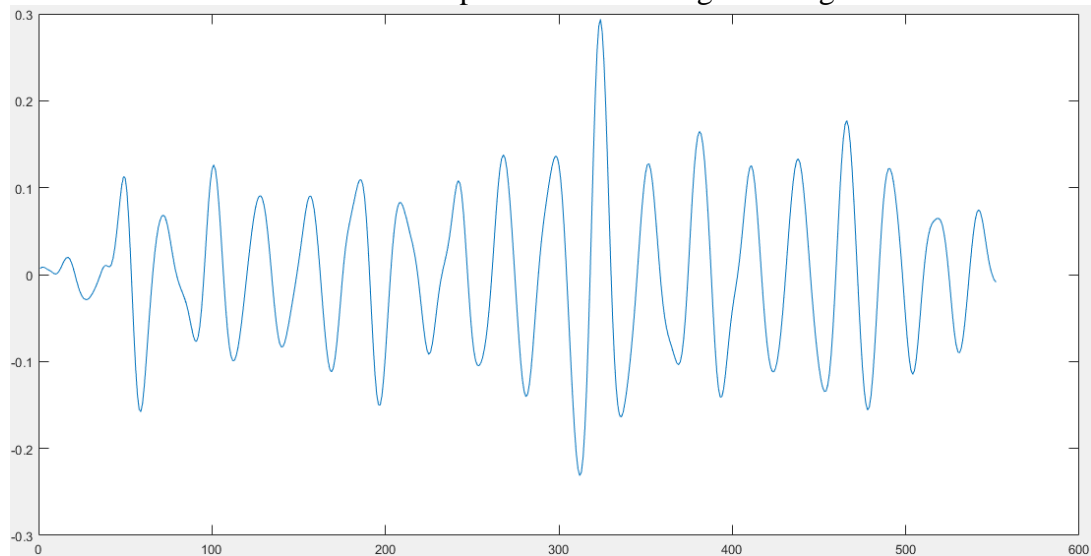


Figura 3.12: Señal rPPG después de suavizar mediante MA.

La señal rPPG final, como se muestra en la figura anterior, posee un marcado carácter periódico y apenas posee ruido. Este es el resultado final del proceso de desenmascaramiento de la señal rPPG que ha ido realizando el algoritmo en los bloques anteriores.

3.4.5 Bloque V: análisis frecuencial de la señal rPPG

Este bloque será el encargado de calcular el pulso cardíaco a partir de la señal rPPG mostrada en la figura 3.12. El método de análisis espectral que se ha seguido es el de la Transformada Rápida de Fourier (FFT), ya que es el más utilizado en estudios anteriores y el que ha dado los resultados más estables. Para ello se ha utilizado la función de Matlab *fft*. El número de puntos que se ha usado para realizar la FFT ha sido de $N=1024$ puntos. Esto permite tener buena resolución en frecuencia y que los escalones entre frecuencias sean más bajos.

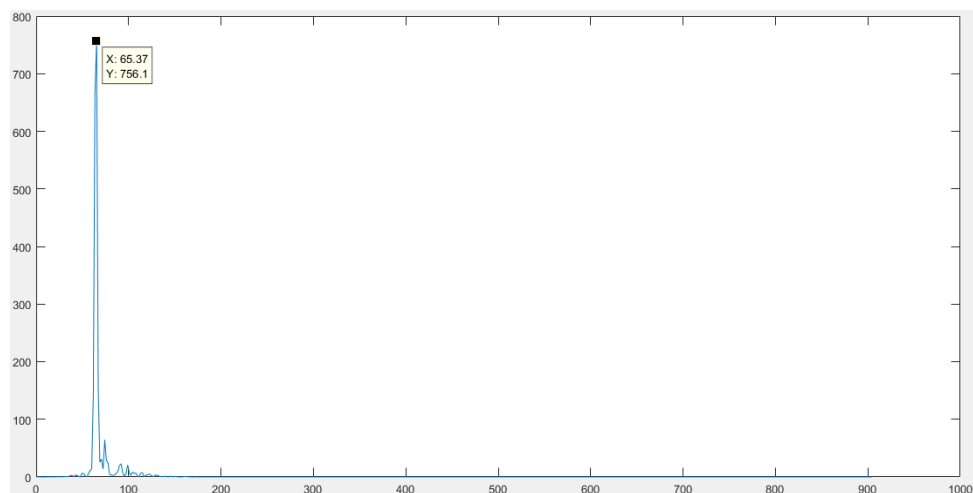


Figura 3.13: Señal rPPG después de suavizar mediante MA.

Como se observa en la Figura anterior, el espectro de esta señal tiene un pico muy marcado correspondiente con el periodo de la señal rPPG, que corresponde con el pulso cardíaco final calculado. Por lo tanto, el pulso cardíaco final corresponderá con aquella frecuencia que marque el pico con mayor energía del espectro.

3.5 Desarrollos no incluidos:

En este apartado se resumirán todos aquellos desarrollos que no han sido incluidos en el algoritmo final, bien porque los resultados no han sido los esperados o bien porque aumentaban el coste computacional drásticamente. Estos desarrollos han sido:

- **Descomposición piramidal de cada *frame*:** Este desarrollo, propuesto por el MIT [8], no se ha incluido debido a que no aportaba ninguna mejora al algoritmo propuesto. La descomposición piramidal sirve para realizar una amplificación de la señal rPPG, de tal manera que se pueda visualizar a simple vista y este no es el objetivo de este algoritmo.
- **Filtrado *Butterworth* de orden 5:** El filtro *Butterworth* de orden 5 propuesto por Unakafov [14] no ha sido incluido en el algoritmo final debido a que no ha dado unos resultados óptimos en todos los vídeos.
- **Utilización de señal *raw* G como la señal rPPG:** La señal de color verde ha sido utilizada en numerosos estudios, como en el Verkrusse et al. [6], como la señal rPPG. Sin embargo, este método no ha dado los mejores resultados debido a que esta señal es muy ruidosa. Además, la señal foto-pletismográfica no solo está presente en la señal de color verde, sino que se ha demostrado en otros estudios, como el de Wang et al. [24], que utilizando una combinación de las señales *raw* para obtener la señal rPPG da como resultado una señal con una periodicidad más reforzada y, por lo tanto, unos mejores resultados.
- **Utilización de ICA como método para obtener la señal rPPG:** Este ha sido descartado porque no está clara cuál de las señales que ICA da como resultado es la que se corresponde con la señal rPPG.
- **Post-procesado de la señal rPPG con un filtro adaptativo *wavelet filter*:** El filtrado adaptativo ha sido eliminado ya que, no solo no aportaba ninguna mejora al algoritmo, sino que en muchas ocasiones daba un resultado erróneo. Se ha comprobado que retirando este filtro aumentaba la precisión del algoritmo.
- **Cálculo del pulso cardíaco mediante conteo de picos de la señal rPPG:** Este método ha sido descartado debido a que la señal rPPG puede tener picos espurios de alta frecuencia causados por el ruido. Esto supondría un resultado erróneo del algoritmo.
- **Cálculo del pulso cardíaco mediante análisis espectral utilizando la Transformada del Coseno:** Esta transformada no ha sido utilizada en el algoritmo final debido a que daba los mismos resultados que la FFT pero con mayor coste computacional.

- **Cálculo del pulso cardíaco instantáneo utilizando la Transformada Wavelet Continua:** El cálculo del pulso cardíaco instantáneo no ha sido incluido en el algoritmo final debido a que aumentaba de manera considerable el coste computacional.

4 Evaluación y resultados

4.1 Introducción

En este capítulo se procederá a evaluar el algoritmo propuesto. Para ello, se han creado tres conjuntos de vídeos y en cada uno de ellos se analizará el resultado del algoritmo propuesto en función de tres parámetros variables: la distancia de análisis, el tipo de iluminación y la resolución. En primer lugar, se explicará el contenido del *dataset* a analizar y el *Ground Truth* utilizado para la verificación de los resultados. Después se evaluarán todos los resultados y, finalmente, se indicarán cuáles son las condiciones óptimas para la ejecución de dicho algoritmo.

4.2 Marco de evaluación

4.2.1 Dataset y Ground Truth

El *Dataset* está formado por vídeos que evalúan el algoritmo en función de tres parámetros: iluminación, resolución y distancia del sujeto. Dentro de cada parámetro se han escogido los siguientes valores:

- **Iluminación:** artificial y natural.
- **Resolución:** 1920x1080, 1280x720 y 640x480.
- **Distancia:** 0,5 metros, 1,5 metros, y 2,5 metros.

Por tanto, los vídeos se dividen en tres grupos diferenciados. Cada uno de estos grupos mide el comportamiento del vídeo en función de la variación de dos parámetros. Estos grupos son:

- **Resolución/Iluminación:** Este grupo está formado por los vídeos que miden el algoritmo con resoluciones e iluminaciones diferentes.
- **Resolución/Distancia:** Estos vídeos miden el algoritmo utilizando diferentes resoluciones y diferentes distancias al sujeto.
- **Distancia/Iluminación:** Estos vídeos miden el algoritmo en función de diferentes distancias e iluminaciones.

Cada uno de estos vídeos tiene una duración aproximada de entre 10 y 20 segundos y han sido grabados mediante la cámara delantera del teléfono móvil Huawei P20-lite.

En todos los vídeos se ha utilizado un *Ground Truth* que evalúe la precisión del resultado del algoritmo. Este *Ground Truth* ha sido el medidor de tensión de la marca OMRON modelo M2. Se ha escogido este modelo debido a su precisión para medir el pulso cardíaco. Por lo tanto, el valor del pulso cardíaco medio de referencia de cada vídeo corresponderá con el valor medido por el *Ground Truth*.



Figura 4. 1: Representación del *Ground Thruth*

4.2.2 Métricas:

Para medir la precisión del algoritmo propuesto, se calculará en cada vídeo el porcentaje de error que hay entre el valor medido por el algoritmo y el valor de referencia dado por el *Ground Thruth*. El porcentaje de error se medirá de la siguiente manera:

$$Error(\%) = \left| \frac{Pulso\ de\ referencia_{Ground\ Thruth} - Pulso\ Medido_{Algoritmo}}{Pulso\ de\ referencia_{Ground\ Thruth}} \right| \times 100$$

Figura 4. 2: Cálculo del porcentaje de error.

Esta fórmula calcula el porcentaje de error que hay entre el pulso medio medido durante la grabación del vídeo por el *Ground Thruth* y el pulso medio medido por el algoritmo propuesto.

4.3 Pruebas y resultados:

En este apartado se realizará la evaluación del algoritmo para los tres conjuntos de vídeos. Para ello, se calculará el porcentaje de error en cada vídeo de cada conjunto de vídeos; por ejemplo: para el conjunto de vídeos resolución/distancia, se analizarán tres tipos de resolución y tres distancias distintas, es decir, serán analizados nueve vídeos. Los resultados serán incluidos en una tabla de tamaño 3 x 3. En total se evaluarán hasta 21 vídeos entre los tres conjuntos para analizar todos estos parámetros.

4.3.1 Resolución/Iluminación

En primer lugar, se evaluará el algoritmo propuesto en función de los parámetros resolución e iluminación. Las resoluciones escogidas para medir el funcionamiento del algoritmo son los tres tipos de resoluciones que ofrece cualquier cámara de bajo coste. Estas son: 1920x1080, 1280x720 y 640x480. La iluminación, por su parte, es un parámetro

importante debido a que la señal rPPG se corresponde con la luz reflejada en la piel del sujeto bajo análisis. Se han tomado como referencia dos tipos de iluminación: la luz natural solar indirecta y la luz artificial de una bombilla hilo incandescente de 60 W. La primera es poco ruidosa y la segunda es más ruidosa debido al parpadeo que produce la corriente eléctrica. No se valora la diferencia entre la cantidad de iluminación ya que este algoritmo necesita iluminación para funcionar. Por ello, la iluminación que se ha utilizado ha sido constante y que permita mantener la zona de análisis iluminada.

Debido a que en este conjunto de vídeos no se contempla la distancia como parámetro, esta será por defecto de 0,5 m.

El porcentaje de error de los resultados obtenidos por el algoritmo en cada vídeo en comparación con el *Ground Truth* son los siguientes:

Resolución Iluminación	1920x1080	1220x720	640x480
Natural	0,95%	3,27%	0,13%
Artificial	0,53%	1,49%	0,96%

Tabla 4. 1: Porcentaje de error en algoritmo propuesto en función de la resolución y la iluminación

Como se puede observar, el porcentaje de error es bastante bajo, por lo que se puede determinar que la iluminación no afecta a cada tipo de resolución y viceversa; es decir, ambos parámetros son independientes del resultado del vídeo

4.3.2 Distancia/Resolución

En segundo lugar, se evaluará el algoritmo en función de los parámetros distancia y resolución. En cuanto a la resolución, se han seleccionado los mismos tipos que en el conjunto anterior, es decir, 1920x1080, 1280x720 y 640x480. Por otra parte, se han seleccionado tres distancias: 0,5 m, 1,5 m y 2,5 m. Entre cada distancia hay una diferencia de 1 m para que la variación sea relevante.

Debido a que en el anterior conjunto analizado los resultados en función de la iluminación no han afectado al resultado final, se han utilizado los dos tipos de iluminación indistintamente.

Resolución Distancia	1920x1080	1220x720	640x480
0,5	0,53%	1,5%	0,96
1,5	1,10%	1,94%	17,98%
2,5	6,34%	3,16%	21,28%

Tabla 4. 2: Porcentaje de error en algoritmo propuesto en función de la distancia y la resolución

Como se observa en los resultados, el error del algoritmo aumenta drásticamente cuando se aumenta la distancia a la vez que se disminuye la resolución. Esto se debe a que el aumento de la distancia provoca que el sensor de la cámara capte la ROI de manera más pequeña, necesitándose así mayor resolución para obtener un buen resultado. Por ello, un aumento de distancia para una misma resolución provoca un aumento en el porcentaje de error. También se observa que para una distancia de 0,5 m la resolución no influye en el resultado. Se puede determinar que para distancias largas se requerirá la mayor resolución posible.

4.3.3 Distancia/Iluminación

Por último, se evaluará el algoritmo en función de los parámetros distancia e iluminación. Se analizarán los tipos iluminación seleccionada en el primero conjunto, es decir, la natural y la artificial, y las distancias evaluadas en el segundo conjunto, es decir, 0,5 m, 1,5m y 2,5m. Se evaluará el cómo varía el resultado del algoritmo al aumentar la distancia en cada tipo de iluminación.

Debido que en el apartado anterior se ha determinado que para la máxima resolución no se ha obtenido un aumento significativo del error al aumentar la distancia, se ha decidido utilizar en este conjunto la resolución máxima, es decir, 1920x1080.

Iluminación Distancia	Artificial	Natural
0,5	0,54%	0,95%
1,5	1,09%	1,10%
2,5	7,35%	6,34%

Tabla 4. 3: Porcentaje de error en algoritmo propuesto en función de la distancia y la iluminación

Como se puede observar en la tabla, el aumento de la distancia influye mucho en el resultado final, indistintamente de la iluminación que se esté dando.

4.4 Comparativa con algoritmos previos

En este apartado se realizará un análisis comparativo del algoritmo propuesto en este trabajo en contraposición con los algoritmos previos de cálculo de pulso cardíaco mediante análisis de secuencias de vídeo en color. Los algoritmos que se han escogido para hacer este análisis comparativo son: *Eulerian video magnification for revealing subtle changes in the world* [8] desarrollado por el MIT en 2012, y el algoritmo propuesto por Unakafov [14], algoritmo en el que se basa el propuesto en este trabajo. El primero de ellos se centra en amplificar la variación de color que sufre la piel con el paso de la sangre y con ello poder calcular el pulso cardíaco; el segundo es un algoritmo centrado en el cálculo del pulso y está formado por una recopilación de los sub-algoritmos que han dado los mejores resultados en cada bloque del algoritmo grande.

Para realizar la comparación de resultados se utilizará el mismo *dataset* que se ha utilizado previamente para la evaluación del algoritmo propuesto, y por lo tanto, se analizará cada algoritmo en función de los tres parámetros de análisis previos: iluminación, distancia y resolución. El cálculo del error por cada vídeo y las tablas ofrecidas seguirán el mismo patrón que en § 4.3.

4.4.1 Resolución/Iluminación

Los resultados del algoritmo de Unakafov [14] son los siguientes:

Resolución Iluminación	1920x1080	1220x720	640x480
Natural	3,09%	10,6%	3,38%
Artificial	0,5%	1,58%	1,15%

Tabla 4. 4: Porcentaje de error del algoritmo de Unakafov [14] en función de la resolución y la iluminación

Como se observa en la tabla anterior, el algoritmo desarrollado por Unakafov es muy sensible al cambio de luz. Para luz artificial da unos resultados excelentes, mientras que para luz natural el error es más alto.

Los resultados del algoritmo desarrollado por el MIT [8] son los siguientes:

Resolución Iluminación	1920x1080	1220x720	640x480
Natural	7,21%	5,01%	13,56%
Artificial	2,7%	4,76%	7,02%

Tabla 4. 5: Porcentaje de error del algoritmo del MIT [8] en función de la resolución y la iluminación

El algoritmo del MIT, como se aprecia en la tabla, es más sensible a la resolución utilizada para grabar los vídeos, pues el error aumenta considerablemente al disminuirla. Esto puede ser debido a que este algoritmo utiliza el método de amplificación-reconstrucción. En este método, por cada *frame* se realiza un análisis espacial mediante pirámides laplacianas o gaussianas, lo que disminuye el tamaño de los *frames* en función de los niveles que se requieran. El MIT recomienda utilizar entre 3 y 6 niveles [8]. Esto hace que para resoluciones bajas se pierda mucha información espacial, lo que hace aumentar el porcentaje de error.

En la siguiente tabla observamos que para luz natural, es sin duda el algoritmo propuesto el que da los mejores resultados, ya que los otros dos algoritmos tienen un porcentaje de error más elevado. Sin embargo, para luz artificial, el algoritmo de Unakafov [14] da, junto con el propuesto, unos resultados óptimos.

Resolución Iluminación		1920x1080	1220x720	640x480
Natural	Propuesto	0,95%	3,27%	0,13%
	Unakafov	3,09%	10,6%	3,38%
	MIT	7,21%	5,01%	13,56%
Artificial	Propuesto	0,53%	1,49%	0,96%
	Unakafov	0,5%	1,58%	1,15%
	MIT	2,7%	4,76%	7,02%

Tabla 4. 6: Comparación de porcentajes de error de los tres algoritmos en función de la resolución y la iluminación

Por último, en la siguiente gráfica se muestra una comparativa de la media de los errores de los tres algoritmos, es decir, el propuesto en este trabajo, el de Unakafov y el del MIT, en cuanto a la resolución y a la iluminación:

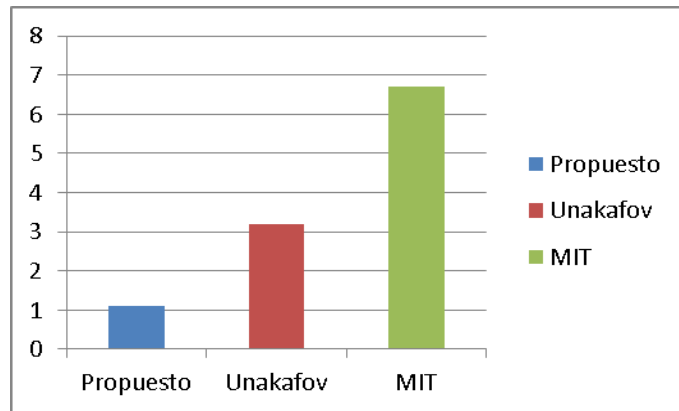


Figura 4. 3: Comparación del porcentaje de error de los tres algoritmos en función de la resolución y la iluminación.

Como se observa, el algoritmo propuesto en este trabajo posee un porcentaje de error muy pequeño en comparación con los otros algoritmos ya que la resolución y el tipo de iluminación afectan mínimamente a este algoritmo.

4.4.2 Distancia/Resolución

A continuación, se muestra la tabla de porcentajes de error del algoritmo de Unakafov [14] en función de la resolución y la distancia:

Resolución \ Distancia	1920x1080	1220x720	640x480
0,5	0,5%	1,58%	1,51%
1,5	0,5%	1,69%	19,64%
2,5	12,7%	3,33%	22,81%

Tabla 4. 7: Porcentaje de error del algoritmo de Unakafov [14] en función de la resolución y la distancia

Como se observa en la tabla anterior, el algoritmo de Unakafov funciona correctamente cuando la ROI detectada tiene un tamaño grande. Sin embargo, el error aumenta considerablemente cuando la distancia de análisis es superior a los dos metros y la resolución es de 640x480. Con esas condiciones, la ROI detectada es muy pequeña por lo que la información obtenida de esta para calcular la señal rPPG será por tanto muy escasa.

La tabla de porcentajes de error para el algoritmo del MIT [8] en función de la resolución y la distancia es la siguiente:

Resolución Distancia	1920x1080	1220x720	640x480
0,5	1,81%	1,58%	1,81%
1,5	3,17%	4,41%	3,57%
2,5	14,28%	16,67%	30,45%

Tabla 4. 8: Porcentaje de error del algoritmo del MIT [8] en función de la resolución y la distancia

La tabla anterior muestra una clara dependencia del algoritmo del MIT con la distancia de análisis, ya que el porcentaje de error aumenta cuando se incrementa esta distancia. Esto puede ser debido a que, al igual que ocurría cuando se disminuía la resolución, al emplear el método de amplificación-reconstrucción, la región de interés se hace más pequeña, haciendo de esta manera que se pierda mucha información para calcular la señal rPPG.

En la tabla inferior se aprecia que el porcentaje de error de los tres algoritmos para la distancia mínima es muy bajo para las tres resoluciones. Sin embargo, según se aumenta la distancia de análisis, se aprecia un crecimiento del error en el algoritmo del MIT [8] para las tres resoluciones por igual, mientras que en los otros dos algoritmos, la distancia de análisis afecta a los 2,5 metros, mayormente para los vídeos con la resolución más baja.

Resolución Distancia		1920x1080	1220x720	640x40
0,5 metros	Propuesto	0,53%	1,5%	0,96
	Unakafov	0,5%	1,58%	1,51%
	MIT	1,81%	1,58%	1,81%
1,5 metros	Propuesto	1,10%	1,94%	17,98%
	Unakafov	0,5%	1,69%	19,64%
	MIT	3,17%	4,41%	3,57%
2,5 metros	Propuesto	6,34%	3,16%	21,28%
	Unakafov	12,7%	3,33%	22,81%
	MIT	14,28%	16,67%	30,45%

Tabla 4. 9: Comparativa de porcentajes de error de los tres algoritmos en función de la distancia y la resolución

En la imagen de abajo se muestra que las medias de los porcentajes de error son altas para los tres algoritmos, ya que estos están entre el 6% y el 8%.

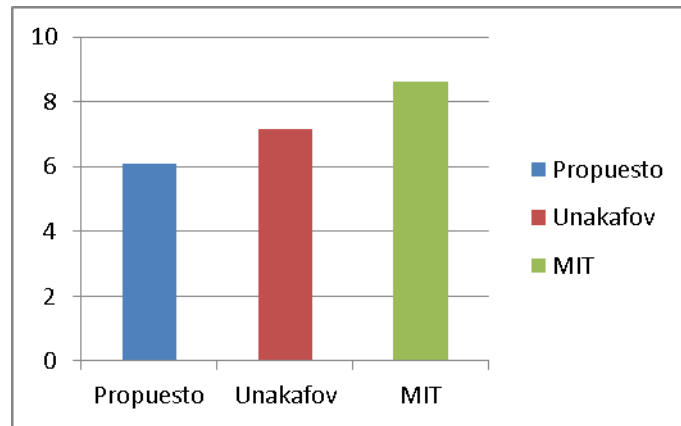


Figura 4. 4: Comparación del porcentaje de error de los tres algoritmos en función de la resolución y la distancia.

Este alto porcentaje de error se debe a que los tres algoritmos son muy sensibles a la distancia de análisis. El algoritmo propuesto es el menos sensible a la distancia y a la resolución utilizada debido a la aplicación de un filtro por el que se descartan todos aquellos píxeles que son detectados como no-piel. Si un píxel es descartado en un solo *frame*, este será descartado en el resto de *frames*, y de esta manera se elimina el ruido producido en la ROI como causa del pequeño tamaño de esta, eliminándose así componentes espurias en frecuencia. Esto no ocurre en el algoritmo de Unakafov [14], puesto que los píxeles detectados como no-piel solo serán descartados en los *frames* donde sean detectados como no-piel, pero no en el resto. Además, al contrario que el algoritmo de Unakafov [14], que utiliza un filtro Butterworth de orden 5, el algoritmo propuesto en este trabajo utiliza un filtro FIR de orden 100. El error del algoritmo del MIT [8] es más alto ya que no realiza ningún tipo de refinamiento de ROI.

Por todo esto, el algoritmo propuesto consigue tener mejor porcentaje de error que el resto.

4.4.3 Distancia/Iluminación

Se procede ahora a realizar la comparación de los tres algoritmos en función de la distancia de análisis y el tipo de iluminación. En primer lugar, en la siguiente tabla se muestra el porcentaje de error del algoritmo de Unakafov.

Iluminación \ Distancia	Artificial	Natural
0,5	0,5%	3,09%
1,5	0,5%	0,5%
2,5	12,7%	12,7%

Tabla 4. 10: Porcentaje de error del algoritmo de Unakafov [14] en función de la distancia y la iluminación

Como se observa, la iluminación natural en el algoritmo de Unakafov hace aumentar el porcentaje de error levemente. Además, como ya se había determinado antes, la distancia de análisis afecta en gran medida al resultado del algoritmo cuando esta es superior a 2,5 metros.

En segundo lugar, se ofrece la tabla con los porcentajes de error del algoritmo del MIT.

Iluminación \ Distancia	Artificial	Natural
0,5	2,7%	7,21%
1,5	3,17%	4,21%
2,5	14,28%	10,34%

Tabla 4. 11: Porcentaje de error del algoritmo del MIT [8] en función de la distancia y la iluminación

En la tabla anterior se aprecia que el algoritmo desarrollado por el MIT es también dependiente del tipo de iluminación, ya que el error aumenta con luz natural. También se aprecia en esta tabla que tiene una gran dependencia de la distancia de análisis.

En la tabla de abajo se observa que los tres algoritmos tienen una gran dependencia con la distancia. Sin embargo, es el algoritmo propuesto el que posee menos dependencia al tipo de iluminación y el que menos aumenta el error con la distancia.

Iluminación \ Distancia		Artificial	Natural
0,5 metros	Propuesto	0,54%	0,95%
	Unakafov	0,5%	3,09%
	MIT	2,7%	7,21%
1,5 metros	Propuesto	1,09%	1,10%
	Unakafov	0,5%	0,5%
	MIT	3,17%	4,21%
2,5 metros	Propuesto	7,35%	6,34%
	Unakafov	12,7%	12,7%
	MIT	14,28%	10,34%

Tabla 4. 12: Comparación de porcentajes de error de los tres algoritmos en función de la distancia y la iluminación

Por último, se muestra a continuación el porcentaje de error en función de la resolución y la distancia en los tres algoritmos:

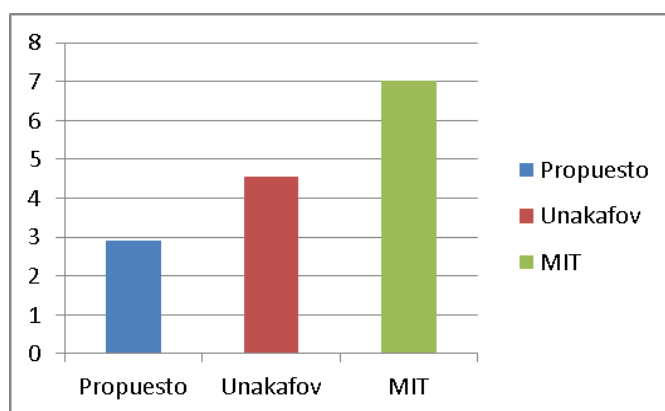


Figura 4. 5: Comparación del porcentaje de error de los tres algoritmos en función de la resolución y la distancia.

La figura anterior muestra que el algoritmo más preciso vuelve a ser el algoritmo propuesto, debido a que este presenta una menor sensibilidad al tipo de iluminación y a la distancia de análisis.

4.5 Conclusiones

En definitiva, como se ha mostrado en este capítulo, el algoritmo propuesto en este trabajo no es igual de sensible a los tres parámetros valorados. Por un lado, partiendo de que la luz es necesaria para el funcionamiento del algoritmo, se ha comprobado que este no es sensible al tipo de luz que hay en el ambiente. Por otro lado, el algoritmo sí es sensible a la distancia y al tipo de resolución. El porcentaje de error aumenta drásticamente cuando la distancia supera 2,5 metros. Este error aumenta todavía más cuando la resolución es menor de 1920x1080. Por lo tanto, para obtener un resultado óptimo del algoritmo, la distancia de análisis no debe superar los 1,5 metros, utilizándose resoluciones mayores a 1220x780 y con un ambiente iluminado, indistintamente del tipo de iluminación que haya.

Las comparaciones realizadas en el apartado anterior muestran que los sub-algoritmos que conforman el algoritmo propuesto dan como resultado que el algoritmo final propuesto en este trabajo de un menor porcentaje de error que los otros dos algoritmos. Estas mejoras implementadas están enfocadas al cálculo claro y preciso de la señal rPPG, ya que esta es la indicadora del pulso. El refinamiento de la ROI es una de las más importantes, ya que evita obtener información temporal de los píxeles que en algún *frame* han sido descartados, y de esta manera se evitan picos espurios en el espectro de la señal rPPG que pueden dar resultados erróneos. Otra mejora implementada que mejora los resultados en el algoritmo propuesto es la aplicación de un filtro ideal paso banda a la ROI para poder descartar aquellos píxeles cuya variación temporal esté fuera de los rangos de interés; en este caso, los que se sitúen fuera del intervalo [50, 180]ppm. Además, la aplicación del filtro FIR de orden 100 en lugar de la aplicación del filtro Butterworth de orden 5 mejora los resultados. No obstante, resulta necesario continuar trabajando en este algoritmo para disminuir todo lo posible la sensibilidad a la distancia de análisis.

5 Conclusiones y trabajo futuro

5.1 Conclusiones

En definitiva, se ha implementado en este trabajo un algoritmo que calcula el pulso cardíaco de una persona mediante la detección de la señal foto-pletismográfica remota (rPPG) a través del análisis de vídeo en color. Estos vídeos deben ser grabados con una cámara convencional, como la cámara de un teléfono móvil. En el análisis de los resultados, el pulso cardíaco detectado en la mayoría de los vídeos del *dataset* ha sido similar o muy parecido al indicado por el *Ground Thruth*. Además, el error mostrado por este algoritmo ha sido muy reducido en comparación con los algoritmos previos del MIT [8] y de Unakafov [14], por lo que la mejora es evidente. Los sub-algoritmos propuestos por Rouast et al. [15] y Unakafov [14] han sido muy relevantes a la hora de implementar el algoritmo final; sin embargo, no todos ellos han dado los resultados esperados. Es por ello que este trabajo también realiza una labor de actualización de los mejores métodos para esta disciplina.

5.2 Trabajo futuro

Después de ver los resultados obtenidos, queda clara la gran dependencia de este algoritmo con la distancia de análisis. Se ha demostrado que para distancias superiores a los dos metros, el porcentaje de error aumenta considerablemente. Es por ello que se propone para trabajos futuros reducir esta dependencia para aumentar la precisión de dicho algoritmo cuando la distancia supera los dos metros. Además de esta mejora, se propone trabajar en:

- Utilizar algún método de detección de pulso cardíaco instantáneo de cara a la utilización de dicho algoritmo en aplicaciones para teléfono móvil. Se propone utilizar la Transformada Wavelet Continua (CWT).
- Implementar un sistema que aplique análisis de componentes principales (PCA) sobre los *frames* del vídeo para obtener de manera más clara la señal rPPG.
- Mejorar el método POS de obtención de la señal rPPG
- Mejorar el post-procesado de la señal rPPG para eliminar componentes periódicas sobrantes. De esta manera, se evitará que aparezcan componentes frecuenciales espurias que posean más energía que la componente frecuencial pulsátil.
- Trabajar sobre otro mapa de color que de mejores resultados.
- Ampliar el dataset con vídeos de diferentes sujetos voluntarios con diferentes colores de piel y, de esta manera, aumentar el conocimiento sobre las condiciones óptimas de ejecución de dicho algoritmo.

6 Referencias

- [1] P. Aragoncillo Ballesteros. «Anatomía del corazón», de *Libro de la salud cardiovascular del Hospital Clínico San Carlos y la Fundación BBVA*, Bilbao, Fundación BBVA, 2009, pp. 35-40.
- [2] M. MacGill y. A. M. García Redondo, «¿Cuál debería ser mi frecuencia cardíaca?» 28/01/2019. [En línea]. Disponible en: <https://www.medicalnewstoday.com/articles/291182.php>. [Última consulta: 30/05/2019].
- [3] R. González Heredia, *El pulso, la auscultación cardíaca y el origen del latido cardíaco: hombres e ideas*, Madrid, Universidad CEU San Pablo, 2013. Tesis dirigida por G. Diéguez Castrillo.
- [4] «Ondas del Electrocardiograma». [En línea]. Disponible en: <https://www.my-ekg.com/generalidades-ekg/ondas-electrocardiograma.html>. [Última consulta: 08/06/2019].
- [5] C. Suero Méndez, «Taller de Arritmias para enfermería, arritmias cardíacas y otros trastornos de conducción». [En línea]. Disponible en: <http://www.hospitalregionaldemalaga.es/LinkClick.aspx?fileticket=yuuUjpcFJfA%3D&tabid=459>. [Última consulta: 08/07/2019].
- [6] W. Verkrusse, L. O. Svaasand, J. Stuart Nelson, «Remote plethysmographic imaging using ambient light», en *Optics Express*, Vol. 16, n.º 26, 2008, pp. 21434-21445. Disponible en: <https://doi.org/10.1364/OE.16.021434>.
- [7] S. M. López Silva, M. L. Dotor, J. P. Silveira, R. Giannetti, L. Herrera. «Fotoplethysmografía por reflexión con LEDs infrarrojos para evaluar órganos y tejidos intra-abdominales: estudio inicial en cerdos», en *Óptica pura y aplicada*, 42 (1), 2009, pp. 23-32.
- [8] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. Freeman, «Eulerian Video Magnification for Revealing Subtle Changes in the World», en *ACM Transactions on Graphics*, Vol. 31, n.º 4, 2012, pp. 1-8.
- [9] G. Balakrishnan, F. Durand, J. Guttag, «Detecting pulse from head motions in video. In Computer Vision and Pattern Recognition (CVPR)», en *Computer Society Conference on Computer Vision and Pattern Recognition*, Portlan, IEEE, 2013, pp. 3430–3437.
- [10] P. Viola, M. Jones, «Rapid object detection using a boosted cascade of simple features», en *Computer Society Conference on Vision and Pattern Recognition*, Vol. 1, Kauai, IEEE, 2001, pp. 511-518.
- [11] J. Shi et al, «Good Features to Track,» *Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [12] B. D. Lucas, T. Kanade, et al., «An iterative image registration technique with an application to stereo vision» in *IJCAI*, vol. 81, pp. 674-679, 1981.
- [13] X. Wu, X. Yang, J. Jin, Z. Yang. «PCA-based magnification method for revealing small signals in video», en *Signal, Image and Video Processing*, Vol. 12, 2018.
- [14] A. M. Unakafov, «Pulse rate estimation using imaging photoplethysmography: generic framework and comparison on a publicly available dataset», en *Biomedical Physics & Engineering Express*, 2018, Vol. 4, Número 4.
- [15] P. V. Rouast, M. T. P. Adam et al. «Remote heart rate measurement using low-cost RGB face video: a technical literature review», en *Springer Nature*, 2018.
- [16] M. Poh, D. J. McDuff, R. W. Picard, «Non-contact, automated cardiac pulse measurements using video imaging and blind source separation». En *Optics Express*, 2010, pp. 10762–10774.

- [17] G. Lempe, S. Zaunseder, T. Wirthgen, S. Zipser, H. Malberg, «ROI selection for remote photoplethysmography». En *Bildverarbeitung für die Medizin*, 2013, pp 99-103.
- [18] X. Li, J. Chen, G. Zhao, M. Pietikäinen, «Remote heart rate measurement from face videos under realistic situations». En *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp 4264–4271.
- [19] D. McDuff, S. Gontarek, R. W. Picard, «Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera». En *IEEE Transactions on Biomedical Engineering*, 2014, pp 2948–2954.
- [20] R. Irani, K. Nasrollahi, T. B. Moeslund, «Improved pulse detection from head motions using dct». En *Proceedings of the 9th International Conference on Computer Vision Theory and Applications*. 2014, pp 118–124.
- [21] B. D. Holton, K. Mannapperuma, P. J. Lesniewski, J. C. Thomas, «Signal recovery in imaging photoplethysmography», en *Physiological Measurement*, 2013, pp 1499–1511
- [22] H. Wu, «Eulerian video processing and medical applications», *Dissertation for the Master Degree*, Cambridge, MA: Massachusetts Institute of Technology, 2012
- [23] D. Rutledge & J-R. Bouveresse, «Independent Components Analysis with the JADE algorithm» *TrAC Trends in Analytical Chemistry*, 2013, 50. 22-32. 10.1016/j.trac.2013.03.013.
- [24] W. Wang, A. den Brinker, S. Stuijk, G. de Haan, «Algorithmic Principles of Remote-PPG». *IEEE Transactions on Biomedical Engineering*, 2016 PP. 10.1109/TBME.2016.2609282.
- [25] G. Balakrishnan, «Analyzing Pulse From Head Motions in Video», Massachusetts Institute Of Technology, 2014.
- [26] G. Androustos, M. Karamanou, C. Stefanadis, «William Harvey (1578-1657): discoverer of blood circulation», *Hellenic J Cardiol*, 2012, vol. 53, no 1, p. 6-9.
- [27] R. E. Dudgeon, (Reed. de 2010), «The Sphygmograph: Its History and Use as an Aid to Diagnosis in Ordinary Practice», Nabu Press, ISBN 9781147502114
- [28] Cuarto Cardal, «El corazón y sus partes». [En línea]. Disponible en: <http://cuarto-cardal.blogspot.com/2007/09/el-corazn-sus-partes.htm>. [Última consulta: 07/06/2019].
- [29] H. V. Carter, H. Gray, «Anatomy of the Human Body», Philadelphia y New York: Lea and Febiger, 1918.